

# SPIM

## Thèse de Doctorat



école doctorale sciences pour l'ingénieur et microtechniques  
UNIVERSITÉ DE TECHNOLOGIE BELFORT-MONTBÉLIARD

N° 2 | 0 | 9

THÈSE présentée par

**Dhouha ATTIA**

pour obtenir le

Grade de Docteur de

l'Université de Technologie de Belfort-Montbéliard

Spécialité : **Informatique**

## Segmentation d'images par combinaison adaptative couleur/texture et classification de pixels

Application à la caractérisation de l'environnement de réception des signaux  
GNSS

Soutenue le 3 octobre 2013 devant le Jury :

P. GOUTON	Président de Jury	Professeur à l'Université de Bourgogne, Dijon (France)
F. DORNAIKA	Rapporteur	Professeur à l'Université du Pays Basque, Leioa (Espagne)
D. HAMAD	Rapporteur	Professeur à l'Université du Littoral Côte d'Opale, Calais (France)
A. BENAZZA	Examinatrice	Professeur à École SUP'COM, Ariana (Tunisie)
S. LEFEBVRE	Examineur	Chargé de projets à FCS Railenium, Famars (France)
J. MARAIS	Co-encadrante	Chargée de recherche à IFSTTAR, Lille (France)
C. MEURIE	Co-encadrant	Chargé de recherche à IFSTTAR, Lille (France)
Y. RUICHEK	Directeur de thèse	Professeur à l'Université Technologie de Belfort-Montbéliard, Belfort (France)



# REMERCIEMENTS

LES travaux de recherche présentés dans ce manuscrit ont été réalisés au sein du laboratoire Systèmes et Transports (SeT) de l'Institut de Recherche sur les Transports, l'Énergie et la Société (IRTES) rattaché à l'Université de Technologie de Belfort-Montbéliard (UTBM). Je tiens pour cela à remercier les directeurs successifs du SeT à savoir M. Abderrafiaa KOUKAM et M. Yassine RUICHEK, pour m'avoir accueilli au sein de leur laboratoire. Je tiens à remercier également le conseil régional de Franche-Comté pour avoir financé ma thèse de doctorat à travers le projet Région ViLoc (coordonné par Cyril MEURIE). Enfin, je remercie indirectement le MEEDDM-DRI pour avoir financé le projet Prédit CAPLOC (coordonné par Juliette MARAIS de l'IFSTTAR et impliquant l'UTBM/IRTES-SeT) et donc m'avoir permis de valider et développer davantage ces travaux de thèse à la problématique de localisation précise de mobiles en environnement urbain.

Je tiens à exprimer ma gratitude à M. Fadi DORNAIKA, Professeur à l'Université du Pays Basque (Espagne) et M. Denis HAMAD, Professeur à l'Université de Littoral Côte d'Opale (France) pour avoir accepté d'être rapporteurs de ma thèse.

Je remercie M. Pierre GOUTON, Professeur à l'Université de Bourgogne (France), pour avoir accepté de présider mon jury.

Je tiens à remercier Mme Amal BENAZZA, Professeur à l'École Supérieure de Communications de Tunis (Tunisie), pour m'avoir fait l'honneur d'accepter de participer à mon jury.

Je remercie M. Sébastien LEFEBVRE, Chargé de Projets à l'Institut de Recherche Technologique RAILENIUM (France), pour avoir accepté, malgré un emploi de temps chargé, de participer à mon jury.

Je remercie sincèrement mon directeur de thèse, M. Yassine RUICHEK, Professeur à l'Université de Technologie de Belfort-Montbéliard (France) et Directeur du laboratoire IRTES-SeT, pour avoir dirigé mon doctorat et guidé mes recherches.

J'adresse mes plus chaleureux remerciements à mes deux co-directeurs de thèse, M. Cyril MEURIE, ancien Maître de Conférences à l'UTBM/IRTES-SeT et maintenant Chargé de Recherche à l'IFSTTAR-LEOST, et Mademoiselle Juliette MARAIS, Chargée de Recherche à l'IFSTTAR-LEOST, pour avoir orienté mon travail à travers leurs remarques et conseils constructifs qui m'ont permis de travailler plus sereinement. En particulier, je tiens à exprimer ma reconnaissance à M. MEURIE, pour m'avoir initialement proposé ce sujet passionnant et pour m'avoir accordé sa confiance dès

le début. Je tiens à lui exprimer toute ma gratitude, pour m'avoir toujours encouragé et encadré sur la partie traitement et analyse des images y compris après son départ à Lille. Je remercie également Mademoiselle MARAIS, pour son encadrement et l'attention qu'elle m'a toujours accordée pour acquérir des connaissances en localisation.

Il m'est très difficile de remercier toutes les personnes qui ont contribué de près ou de loin à ce travail. Je veux néanmoins remercier Frédéric ZANN, Ingénieur d'Études à l'UTBM/IRTES-SeT et Amaury FLANQUART, Ingénieur d'Études à l'IFSTTAR-LEOST pour le temps non négligeable qu'ils ont passé pour m'aider à manipuler les plateformes d'acquisitions d'images et de données GPS.

J'adresse également dans une plus large mesure mes remerciements à toutes les personnes de l'équipe : Cindy CAPPELLE, Maître de Conférences à l'UTBM/IRTES-SeT, mes collègues et amis doctorants Lijun, Houssam, Julien, Julian, Madeleine, Baudoin, You et Yang, du laboratoire SeT, dont l'ambiance a donné une autre dimension à ce travail. Un merci tout particulier à Kahina et Ferdaous mes chères amies et sœurs qui m'ont toujours encouragée et soutenue. Je n'oublie pas mes amies Safa et Amani qui malgré la distance et leurs propres soucis ont toujours pensé à moi.

Enfin, j'adresse mes plus affectueux remerciements à mes parents, mes sœurs Afef et Nesrine, mon petit frère Mohamed et à toute ma famille pour leur précieux soutien et leurs encouragements permanents. Merci à mes parents qui ont cru en moi et attendent depuis plusieurs années mon retour sans cesser de prier pour moi.



*à maman et papa : source de tendresse inépuisable ; l'origine*

*à mon frère et mes sœurs : soutien à vie, le présent*

*à toi mon âme sœur, l'avenir*

*Dhouha*

إلى أمي و أبي  
إلى أخي و أخواتي  
إليكم جميعا أهدي عملي التواضع هذا

ضمي



# SOMMAIRE

<b>Introduction générale</b>	<b>15</b>
<b>1 Segmentation d'images en utilisant la couleur et la texture</b>	<b>19</b>
1.1 Introduction	19
1.2 La couleur, la texture	20
1.2.1 La couleur	20
1.2.1.1 La définition physiologique de la couleur	20
1.2.1.2 Les modèles de représentation de la couleur	20
1.2.1.3 La couleur dans l'image	21
1.2.2 La texture	22
1.2.2.1 Attributs géométriques	25
1.2.2.2 Attributs basés sur la modélisation spatiale	25
1.2.2.3 Attributs spatio-fréquentiels	29
1.2.2.4 Attributs statistiques	29
1.3 Segmentation d'images par utilisation conjointe de la couleur et de la texture	34
1.3.1 Intégration implicite	35
1.3.1.1 Compression-based Texture Merging (CTM)	36
1.3.1.2 MeanShift	36
1.3.1.3 Normalized Cuts (NCuts)	38
1.3.2 L'intégration successive	40
1.3.2.1 Jseg	40
1.3.2.2 Fusion Clustering Results (FCR)	41
1.3.2.3 Texture and Boundary Encoding-based Segmentation (TBES)	43
1.3.2.4 Segmentation de Felz-Hutt (FH)	45
1.3.3 L'intégration lors du processus de segmentation	46
1.3.3.1 Segmentation par une fonction de variation totale (Total Variation segmentation (TV))	47
1.3.3.2 Segmentation basée sur le modèle Chan-Vese (CV)	48

1.3.3.3	Segmentation basée sur la cohérence couleur/texture (C-Tex) . . . . .	48
1.3.3.4	Segmentation par combinaison de luminance, couleur et texture (CTB) . . . . .	49
1.3.3.5	Segmentation par combinaison couleur/texture fixe (CF) . . . . .	51
1.3.3.6	Segmentation par combinaison couleur/texture adaptative ou supremum (CA, CS) . . . . .	52
1.4	Conclusion . . . . .	53
<b>2</b>	<b>Évaluation de la segmentation et bases de test</b>	<b>57</b>
2.1	Introduction . . . . .	57
2.2	Les méthodes d'évaluation de la segmentation . . . . .	58
2.2.1	Les différentes classifications des méthodes d'évaluation . . . . .	58
2.2.2	Méthodes d'évaluation supervisée . . . . .	61
2.2.2.1	La mesure de RAND . . . . .	63
2.2.2.2	La mesure de JACCARD . . . . .	64
2.2.2.3	La mesure de FOWLKES et MALLOWS . . . . .	64
2.2.2.4	La mesure d' <i>information mutuelle</i> . . . . .	65
2.2.2.5	La mesure de <i>variation d'information</i> . . . . .	65
2.2.2.6	La mesure de VAN DONGEN . . . . .	66
2.2.2.7	La mesure de <i>Bipartite Graph Matching</i> . . . . .	66
2.2.2.8	La mesure de MARTIN . . . . .	67
2.2.2.9	La mesure de VINET . . . . .	68
2.2.2.10	La mesure de YASNOFF . . . . .	68
2.2.2.11	La mesure de HUANG-DOM . . . . .	69
2.2.2.12	Autres métriques d'évaluation . . . . .	69
2.2.3	Méthodes d'évaluation non-supervisée . . . . .	70
2.2.3.1	La mesure de ZEBOUDJ . . . . .	71
2.2.3.2	Les mesures de LEVINE et NAZIF . . . . .	72
2.2.3.3	La mesure de dissimilarité de LIU et YANG . . . . .	73
2.2.3.4	La mesure de BORSOTTI . . . . .	73
2.2.3.5	L'erreur quadratique moyenne . . . . .	74
2.2.3.6	Autres métriques d'évaluation non-supervisée . . . . .	74
2.3	Les bases de tests/apprentissage/validation . . . . .	74
2.3.1	Bases d'images couleur . . . . .	75
2.3.1.1	La base d'images de MCGILL . . . . .	75

2.3.1.2	La base d'images de BERKELEY . . . . .	75
2.3.1.3	La base d'images de PASCAL VOC . . . . .	77
2.3.2	Bases d'images texture . . . . .	77
2.3.2.1	La base d'images de VISTEX . . . . .	77
2.3.2.2	La base d'images de OUTEX . . . . .	78
2.3.2.3	La base d'images de PRAGUE . . . . .	78
2.3.2.4	La base d'images de la librairie de texture . . . . .	79
2.4	Comparatif des méthodes de segmentation couleur/texture . . . . .	79
2.4.1	Choix de bases de tests . . . . .	79
2.4.2	Choix et développement des métriques d'évaluation . . . . .	80
2.4.3	Expérimentation et discussions . . . . .	82
2.5	Conclusion . . . . .	96
<b>3</b>	<b>Une méthode de combinaison couleur-texture non paramétrique</b>	<b>101</b>
3.1	Introduction . . . . .	101
3.2	Les descripteurs de couleur et de texture . . . . .	102
3.2.1	Les descripteurs de couleur . . . . .	103
3.2.1.1	Gradient morphologique couleur . . . . .	103
3.2.1.2	Gradient de DI-ZENZO . . . . .	103
3.2.2	Les descripteurs de texture . . . . .	105
3.2.2.1	Gradient basé sur les filtres de Gabor . . . . .	105
3.2.2.2	Gradients basés sur les paramètres d'HARALICK . . . . .	106
3.2.2.3	Gradient basé sur le filtre alterné séquentiel . . . . .	111
3.3	Une nouvelle méthode non paramétrique de combinaison couleur/texture . . . . .	117
3.3.1	L'analyse en composantes principales . . . . .	117
3.3.2	La combinaison couleur/texture par comparaison de valeurs propres . . . . .	120
3.4	La méthode de Ligne de Partage des Eaux . . . . .	124
3.5	Résultats expérimentaux . . . . .	127
3.6	Conclusion . . . . .	140
<b>4</b>	<b>Caractérisation de l'environnement de réception des signaux GNSS</b>	<b>141</b>
4.1	Introduction . . . . .	141
4.2	État de l'art . . . . .	142
4.3	Les projets ViLoc, CAPLOC . . . . .	144
4.4	Méthode d'évaluation de la caractérisation de l'environnement . . . . .	147
4.5	Stratégie de détection du ciel . . . . .	148

4.5.1	Première approche : Segmentation par combinaison couleur/texture et classification de régions . . . . .	148
4.5.2	Deuxième approche : Simplification de l'image et classification pixellaire . . . . .	154
4.5.2.1	Filtre médian . . . . .	156
4.5.2.2	Filtre moyenneur . . . . .	156
4.5.2.3	Filtre gaussien . . . . .	157
4.5.2.4	Filtre exponentiel symétrique . . . . .	157
4.5.2.5	Filtre de MALIK-PERONNA . . . . .	158
4.5.2.6	Filtre de NAGAO . . . . .	159
4.5.2.7	Filtre adaptatif basé sur le voisinage . . . . .	159
4.5.2.8	Filtre adaptatif Symetric Nearest Neighbourhood (SNN) . . . . .	160
4.5.2.9	Reconstruction Géodésique par Dilatation (RGD) . . . . .	161
4.5.3	Classification pixellaire . . . . .	165
4.5.3.1	Méthodes de classification non-supervisée . . . . .	165
4.5.3.2	Méthodes de classification supervisée . . . . .	168
4.6	Paramétrage de la seconde approche proposée . . . . .	170
4.7	Conclusion . . . . .	180
<b>5</b>	<b>Utiliser l'image pour localiser avec plus de précision</b>	<b>181</b>
5.1	Introduction . . . . .	181
5.2	Introduction à la localisation par satellites . . . . .	182
5.2.1	Le système GNSS . . . . .	183
5.2.2	Les données de mesure . . . . .	183
5.2.3	Le calcul de la position . . . . .	183
5.3	Impact de l'environnement de réception sur la précision de la localisation . . . . .	184
5.3.1	Définition du DOP . . . . .	184
5.3.2	L'imprécision de la mesure de pseudo-distance causée par les multi-trajets . . . . .	185
5.3.3	Méthodes de mitigation du problème de multi-trajet . . . . .	186
5.4	Application de techniques de traitement d'images pour la localisation . . . . .	187
5.4.1	Repositionnement des satellites dans l'image . . . . .	188
5.4.2	Détection de la ligne d'horizon . . . . .	190
5.5	Calcul d'une nouvelle position . . . . .	191
5.5.1	Systèmes d'acquisition . . . . .	191
5.5.2	Description des données expérimentales . . . . .	191

5.5.3	Calcul de position à partir des satellites LOS uniquement . . . . .	192
5.5.4	Utilisation combinée du DOP et de l'exclusion des satellites NLOS .	196
5.5.5	Discussions sur l'impact de la végétation sur la visibilité satellitaire .	198
5.5.6	Analyse quantitative . . . . .	200
5.5.6.1	Estimation de la précision d'une position calculée . . . . .	200
5.6	Conclusion . . . . .	204
<b>Conclusion générale</b>		<b>207</b>
<b>Bibliographie</b>		<b>229</b>
<b>Liste des figures</b>		<b>238</b>
<b>Liste des tables</b>		<b>241</b>





# INTRODUCTION GÉNÉRALE

La segmentation d'images est une étape cruciale pour plusieurs applications de la vision par ordinateur. Elle a reçu une forte attention aussi bien de la part des chercheurs académiques que des industriels durant les trois dernières décennies du  $XX^{\text{ème}}$  siècle. En effet, nous faisons souvent appel à une étape de segmentation d'images pour des applications de localisation, d'imagerie médicale, de vidéo surveillance, de télédétection, d'indexation multimédia ou encore de compression d'images, etc. Malgré l'évolution rapide des performances des méthodes de segmentation, nous ne pouvons pas encore parler d'une méthode générique s'appliquant à différents types d'images et prenant en compte tous types d'applications. Une méthode de segmentation d'images est généralement conçue et adaptée aux objectifs de l'application à laquelle elle est dédiée. La difficulté de trouver une méthode de segmentation générique peut être liée à la difficulté que trouve la communauté en vision par ordinateur à définir la segmentation d'images elle-même. En effet, la définition de la segmentation d'images renferme implicitement de l'ambiguïté. Bien que la plupart des auteurs s'accordent sur le fait que la segmentation est un processus de partitionnement de l'ensemble des pixels d'une image en plusieurs sous-ensembles disjoints selon des critères prédéfinis, ils divergent sur le choix de ces critères qui restent généralement une source de débat. Un premier groupe d'auteurs considère la notion d'homogénéité et définit la segmentation d'images comme une partition de l'image en différentes régions telles que chacune d'elles (et non l'union de deux régions) soit homogène [Cheng 2001]. De même, MORRIS et al. [Morris 1986] décrivent la segmentation d'images comme un processus de partitionnement en régions homogènes mais différentes des régions voisines. SKARBEEK et KOSCHAN [Skarbek 1994] optent plutôt pour une définition plus simple qui consiste seulement à détecter les régions homogènes. Un deuxième groupe d'auteurs part sur le principe qu'une méthode de segmentation d'images est l'imitation de l'interprétation humaine de la scène analysée. Parmi les auteurs qui adoptent cette définition, FU et MUI [Fu 1981] considèrent le problème de segmentation comme une sorte de perception psychophysique qui rend la résolution du problème inenvisageable par des solutions analytiques. MARTIN et al. [Martin 2001] rejoignent le point de vue de FU et MUI en comparant le résultat de segmentation automatique à des segmentations de référence humaines. Un troisième groupe d'auteurs admet les deux précédents points de vue [Shi 2000, Nock 2004a, Adamek 2005], ce qui à notre sens montre la multitude de visions et donc la difficulté de définir le problème de segmentation d'images.

Quelle qu'elle soit, la segmentation d'images constitue souvent une étape intégrante de tout processus de traitement d'images. Par conséquent, l'algorithme de segmentation utilisé influence directement la performance du système global. Le choix de la méthode de segmentation est fonction de l'application visée et de ce fait, il constitue une condition essentielle pour garantir un fonctionnement performant du système.

Certaines applications nécessitent une simple division de l'image en régions grossières et homogènes alors que d'autres nécessitent une détection plus précise et sémantique de la scène. D'autre part, certains systèmes exigent une grande précision alors que d'autres requièrent un traitement temps réel. Par conséquent, tous ces critères agissent sur le fonctionnement de l'algorithme de segmentation. La couleur et la texture font partie des sources d'information permettant de décrire les objets d'une image. Elles sont utilisées de différentes manières dans les algorithmes de segmentation. Certains auteurs utilisent séparément la couleur et la texture mais les résultats obtenus ne sont généralement pas suffisamment satisfaisants (ou du moins peuvent être améliorés). Plusieurs travaux ont prouvé qu'une utilisation conjointe de ces deux sources d'information (la couleur et la texture), fortement liée à la perception humaine, permet de proposer de nouvelles méthodes de segmentation plus robustes et précises dans la description du contenu de l'image. La première partie des travaux réalisés dans cette thèse s'inscrit dans cette logique. Nous présentons une méthode de combinaison adaptative des informations de couleur et de texture, utilisée ensuite en entrée d'un algorithme de segmentation morphologique couleur (la méthode des Lignes de Partage des Eaux) permettant d'offrir des résultats très satisfaisants comparés aux méthodes de segmentation connues de la littérature.

## PROJETS ET APPLICATIONS

Comme indiqué ci-dessus, la segmentation d'images ou plus généralement le traitement d'images est utilisé dans un certain nombre d'applications et notamment pour l'amélioration de la précision de la localisation d'un mobile en milieu urbain. Ceci concerne la deuxième partie de cette thèse où nous avons cherché à caractériser l'environnement de réception des signaux GNSS par traitement d'images. Ce travail s'inscrit dans le cadre de deux projets coopératifs entre l'IRTES-SET (UTBM) et l'IFSTTAR-LEOST (le projet ViLoc financé par la région Franche-Comté et le projet CAPLOC financé par le PREDIT et coordonné par l'IFSTTAR-LEOST). Dans ce contexte, la localisation d'un mobile, fonction clé des applications relevant des Systèmes de Transport Intelligent (ITS), est généralement remplie par un récepteur GNSS (le GPS par exemple). Celle-ci s'avère très souvent efficace. Cependant, les applications à caractère sécuritaire imposent des contraintes plus fortes en termes d'intégrité et de précision de l'information de position qui ne sont pas assurées par les GNSS dans les environnements fortement contraints tels que les zones urbaines denses. En effet, dans ce type d'environnements, les signaux peuvent être réfléchis voire complètement bloqués par des obstacles tels que les bâtiments, la végétation, les tunnels, etc. Ceci entraîne une erreur dans l'estimation du temps et donc une imprécision dans la localisation. L'arrivée de GALILEO, bien qu'améliorant la précision de localisation, ne permet pas de s'affranchir des phénomènes de multi-trajets présents en milieu urbain dense et qui rendent la mesure de positionnement imprécise. Pour essayer de s'affranchir de ces limitations, nous avons, dans la deuxième partie de cette thèse, développé une stratégie permettant la caractérisation de l'environnement de réception des signaux GNSS en se basant sur la fusion des informations provenant des systèmes GNSS et de vision.

Dans cette thèse, nous avons proposé deux approches permettant de détecter dans l'image, en temps réel, la zone de " ciel " (où les signaux sont reçus en direct) et de "

non ciel " (les signaux sont reçus par réflexions ou totalement bloqués) afin de pouvoir, par la suite, repositionner des satellites dans l'image, choisir pour le calcul de la position les satellites estimés "sans erreur de mesure". La première stratégie consiste à appliquer une segmentation basée sur la combinaison couleur/texteure proposée et développée dans la première partie de cette thèse, suivie d'une classification de régions. La deuxième stratégie consiste à simplifier l'image initiale par une opération de morphologie mathématique puis à appliquer une classification pixellaire adaptée. Afin de renforcer l'objectif visé, nous avons comparé deux techniques d'estimation de positions tenant compte du résultat de traitement d'images à partir d'un même jeu de signaux satellitaires reçus/images enregistrées. Cette comparaison nous a conduit donc à déterminer le choix approprié qui permet un usage optimisé des satellites disponibles pour fournir une trajectoire la plus précise possible.

## PLAN DU MANUSCRIT

Cette thèse s'articule autour de deux volets : un volet théorique (correspondant aux chapitres 1, 2 et 3) et un volet applicatif (correspondant aux chapitres 4 et 5). Notre contribution porte principalement sur la proposition et le développement d'une méthode de segmentation basée sur une combinaison adaptative et non paramétrique de la couleur et de la texture, ainsi qu'une stratégie de caractérisation temps réel de l'environnement de réception des signaux satellitaires basée sur une simplification/segmentation de l'image suivie d'une classification pixellaire adaptée.

Dans le chapitre 1 de ce manuscrit, nous rappelons brièvement les définitions et les propriétés de la couleur et de la texture. Nous effectuons également un état de l'art exhaustif des différents descripteurs de couleur et de texture en soulignant les principaux attributs utilisés dans la littérature. Nous insisterons davantage dans ce premier chapitre sur les méthodes de segmentation utilisant conjointement ces types de descripteurs de couleur et de texture.

Le chapitre 2 présente un état de l'art exhaustif des différentes métriques d'évaluation des méthodes de segmentation d'images utilisées dans la littérature. Deux familles de méthodes d'évaluation sont présentées : l'évaluation supervisée qui compare la segmentation automatique à une/plusieurs segmentation(s) de référence prédéfinies par des experts et l'évaluation non supervisée qui mesure la qualité générale d'une segmentation automatique sans avoir recours à une segmentation de référence.

Dans le chapitre 3, nous présentons une nouvelle approche de combinaison adaptative et non paramétrique des informations de couleur et de texture présentes dans l'image. Dans cette nouvelle combinaison, nous avons eu recours à l'analyse spectrale et à l'étude des axes principaux d'un nuage de points 3D généré à partir d'un ensemble d'attributs couleur et d'attributs texture. L'apport de l'approche proposée consiste à retenir les informations de couleur et de texture les plus pertinentes, permettant d'assurer une segmentation plus précise, à partir d'un ensemble de gradients couleur (le gradient morphologique couleur et le gradient de DI-ZENZO) et de gradients texture (les gradients basés sur trois attributs d'HARALICK, le gradient basé sur le filtre Alternatif Séquentiel et

le gradient basé sur les filtres de GABOR).

Le chapitre 4, premier chapitre du volet applicatif de cette thèse, consiste à caractériser l'environnement de réception des signaux GNSS en milieu urbain. Deux solutions sont présentées. La première consiste à segmenter l'image acquise en mobilité à l'aide de la méthode de segmentation basée sur la nouvelle combinaison couleur/texture proposée, puis à effectuer une classification via un classifieur adapté (sélectionné parmi les principaux classifieurs supervisés et non supervisés de la littérature) afin de classer les pixels de l'image en deux classes " ciel " et " non ciel ". La deuxième solution est basée sur une simplification de l'image par un filtre adapté, suivie de la même étape de classification pixellaire appliquée directement sur l'image simplifiée. Ces deux approches sont testées et évaluées afin de montrer les avantages et les limites de chacune d'elles.

Dans le dernier chapitre du manuscrit (chapitre 5), nous présentons comment cette nouvelle stratégie de caractérisation de l'environnement de réception des signaux GNSS, basée sur l'analyse d'images, permet d'estimer la position d'un véhicule en mouvement dans un milieu urbain dense de manière plus précise. Les résultats obtenus en sortie de la procédure de traitement d'images nous permettent d'éliminer du calcul de la position les signaux qui dégraderaient cette information, pour ne considérer que les signaux utiles. Nous distinguons pour cela deux types de signaux reçus : les satellites situés dans la zone " ciel " dont le signal est reçu via un trajet direct (LOS - Line-Of-Sight), et les satellites situés dans la zone " non ciel " et dont le signal est bloqué ou reçu avec une atténuation trop forte pour être exploité (NLOS - Non-line-Of-Sight). Enfin, nous comparons deux techniques d'estimation de positions à partir d'un même jeu de signaux satellitaires reçus par notre plateforme expérimentale, en considérant également l'information relative à la répartition géométrique des satellites.

Le manuscrit se termine par une conclusion générale dans laquelle nous résumons nos contributions et présentons une synthèse des résultats obtenus. Nous présentons également les voies de recherche et d'amélioration qui s'ouvrent à l'issue de ce travail.

# SEGMENTATION D'IMAGES EN UTILISANT LA COULEUR ET LA TEXTURE

## Sommaire

<b>1.1</b>	<b>Introduction</b>	<b>19</b>
<b>1.2</b>	<b>La couleur, la texture</b>	<b>20</b>
1.2.1	La couleur	20
1.2.2	La texture	22
<b>1.3</b>	<b>Segmentation d'images par utilisation conjointe de la couleur et de la texture</b>	<b>34</b>
1.3.1	Intégration implicite	35
1.3.2	L'intégration successive	40
1.3.3	L'intégration lors du processus de segmentation	46
<b>1.4</b>	<b>Conclusion</b>	<b>53</b>

## 1.1/ INTRODUCTION

La couleur et la texture constituent les deux sources d'information les plus utilisées dans plusieurs applications d'analyse d'images telles que la reconnaissance d'objets, l'indexation multimédia, la compression, la segmentation d'image, etc. Dans ce premier chapitre, nous commençons donc par présenter les définitions relatives à ces deux notions. Nous présentons ensuite un état de l'art exhaustif sur les diverses familles d'attributs de texture existantes dans la littérature tout en soulignant leurs caractéristiques et détaillant leur mode de calcul. En effet, ces attributs sont classés en fonction de leur technique de calcul ainsi que les outils mathématiques mis en œuvre. Dans la deuxième partie de ce chapitre, nous accordons plus d'attention aux méthodes de segmentation utilisant conjointement la couleur et la texture. Nous rappelons donc un ensemble de techniques de segmentation basées sur cette stratégie (utilisation conjointe de la couleur et de la texture) constituant un benchmark dans la littérature. Nous utilisons ces méthodes comme base de référence pour évaluer les performances de la méthode de combinaison couleur/texture qui sera proposée dans le chapitre 3.

## 1.2/ LA COULEUR, LA TEXTURE

Avant d'examiner l'impact de l'utilisation conjointe de la couleur et de la texture dans les méthodes de segmentation, il s'avère intéressant de rappeler les définitions et les caractéristiques de chacune de ces deux sources d'informations.

### 1.2.1/ LA COULEUR

Nous commençons par rappeler brièvement la définition physiologique de la couleur, ce qui permettra de comprendre l'interprétation de ce phénomène par le cerveau humain. Ensuite, nous introduisons, avec plus de détails, les différentes représentations de la couleur. Finalement, nous rappelons le codage de cette source d'information dans une image.

#### 1.2.1.1/ LA DÉFINITION PHYSIOLOGIQUE DE LA COULEUR

La couleur n'est ni une matière, ni une lumière mais plutôt une sensation reçue par l'intermédiaire de l'œil humain et traduite par notre cerveau lors de la vision d'un objet coloré. Par conséquent, la couleur est une sensation physiologique qui dépend d'un ensemble de facteurs :

- la nature de l'objet perçu ;
- la lumière incidente qui permet à l'œil de recevoir le message ;
- l'œil qui perçoit cette lumière et communique cette sensation au cerveau.

#### 1.2.1.2/ LES MODÈLES DE REPRÉSENTATION DE LA COULEUR

La couleur joue un rôle fondamental dans la distinction des objets constituant une scène/image étant donné l'analogie entre le fonctionnement de l'œil humain et les modèles de la vision humaine. Avant d'introduire chacun de ces modèles, nous rappelons que la théorie tri-chromatique de la lumière a été introduite par Thomas Young en 1801 [Pire ] et complétée dans les travaux de HERMANN VON HELMOLTZ en 1804 [Pire ]. Cette théorie précise que la rétine possède trois cônes photorécepteurs couleur sensibles respectivement au bleu, vert et rouge (cf. figure 1.1(a)). Ces éléments de réception, situés dans la rétine, reçoivent les radiations de la lumière incidente dont les longueurs d'onde doivent appartenir au spectre visible avec des longueurs d'onde comprises entre 380 nm et 780 nm et qui répondent chacun selon sa courbe de sensibilité spectrale, notée :

- *S* pour les *Short wavelength* ou bien *B* (cône sensible au bleu).
- *M Medium wavelength* ou bien *G* (cône sensible au vert).
- *L Long wavelength* ou bien *R* (cône sensible au rouge).

Comme le montre la figure 1.1(b), ces courbes sont caractérisées par des pics appelés "pics d'absorption", qui correspondent respectivement aux longueurs d'onde 440 nm, 545 nm et 570 nm.

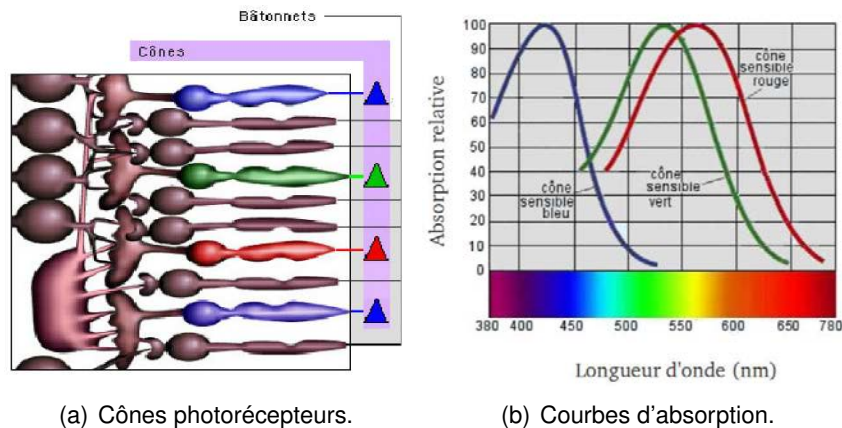


FIGURE 1.1 – Sensibilité spectrale des trois types de cônes photorécepteurs couleur de l'œil humain et des bâtonnets (source : <http://tpelavision.chez-alice.fr/oeil.htm>).

Dans ce qui précède, le modèle *tri-stimulus* a été introduit. Il s'agit du modèle le plus répandu pour le codage des couleurs dans différents espaces. Il existe d'autres modèles de représentation de la couleur dont le principe est rappelé brièvement dans ce qui suit :

- **le modèle des couleurs opposées**, développé par HERING, consiste à considérer la couleur comme une sensation psychologique. HERING s'éloigne des notions physiques et affirme qu'il n'y a pas trois couleurs fondamentales, mais plutôt quatre primaires correspondantes à des sensations psychologiques et qui sont le rouge, le vert, le bleu et le jaune. D'après le fondateur de ce modèle, le jaune peut avoir une dominante rouge ou une dominante verte, mais en aucun cas une dominante bleue en justifiant sa théorie par les couleurs opposées rouge-vert et bleu-jaune. En suivant la même logique, HERING considère la luminosité comme une opposition blanc-noir [Hering 1880, Hurvich 1957] ;
- **le modèle de FAUGERAS**, introduit par ce dernier, consiste à transformer la réponse non-linéaire des cônes photorécepteurs en une réponse proportionnelle au logarithme de la lumière incidente. Les composantes couleur fournies par ce modèle sont deux composantes chromatiques et une troisième composante achromatique [Faugeras 1979] ;
- **Le modèle de PRATT**, instauré par ce dernier dans ses travaux [Pratt 1978], est définit contrairement aux autres modèles, par trois nouvelles notions : la teinte, la saturation et la luminosité.

Les chercheurs ont instauré de nouveaux modèles représentant la couleur dont les plus récents ont été proposés en 1998 et 2002 par la CIE [CIE 1998, CIE 2002]. Dans ce qui suit, nous adoptons le modèle couleur le plus répandu : le modèle tri-chromatique.

#### 1.2.1.3/ LA COULEUR DANS L'IMAGE

Dans le domaine du traitement d'images, une image couleur donnée  $I$  est définie comme étant un ensemble de pixels ayant chacun trois composantes rouge, verte et bleue suite à la numérisation de la scène par une caméra couleur. Ces composantes peuvent être aussi exprimées dans d'autres espaces couleur. Pour de plus amples informations sur ces différents espaces couleur, le lecteur pourra consulter la thèse de VANDENBROUCKE [Vandenbroucke 2000] qui décrit en détails cet aspect.



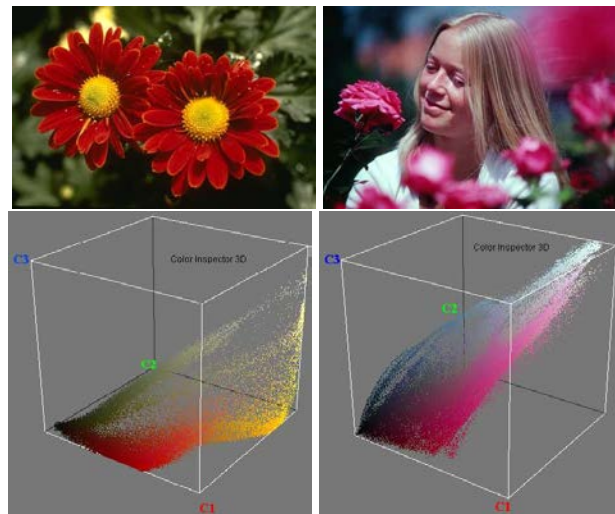


FIGURE 1.2 – Représentation 3D de deux images couleur dans un espace muni du repère  $(C_1, C_2, C_3)$  à l'aide du logiciel "Color Inspector 3D" (source : <http://www.f4.fhtw-berlin.de/~barthel/ImageJ/ColorInspector//help.htm>,).

D'une manière générale, quel que soit l'espace couleur considéré, un pixel  $p$  d'une image est décrit par trois valeurs  $C_1(p)$ ,  $C_2(p)$  et  $C_3(p)$ . Ceci permet de représenter le pixel par un point dans l'espace et de percevoir l'image comme un nuage de points (cf. figure 1.2).

### 1.2.2/ LA TEXTURE

Au même titre que la couleur, la texture est un facteur fondamental dans la perception de l'environnement et la reconnaissance de ses objets. Contrairement à la couleur, la texture reste difficile à définir de manière précise et générique. Le nombre de définitions possibles données dans la littérature en témoignent [Sklansky 1978, Tamura 1978, Haralick 1979]. La définition littéraire de la texture est la "répétition spatiale d'un même motif dans différentes directions de l'espace". Cependant, cette explication est parfois considérée insuffisante étant donné qu'elle est indépendante du comportement de l'observateur humain. À titre d'exemple, la texture d'un tissu ou d'un mur de briques (cf. figure 1.3) peut satisfaire cette définition, ce qui n'est pas toujours le cas pour d'autres textures. La texture possède plusieurs aspects visuels discriminants à savoir l'aspect grossier, fin, granuleux, régulier ou irrégulier (cf. figure 1.4). Par conséquent, l'homogénéité de la surface est le critère à considérer pour différencier deux objets de textures différentes.



FIGURE 1.3 – Exemples de textures répétitives prises de la base de texture VISTEX [VisTex 1995].



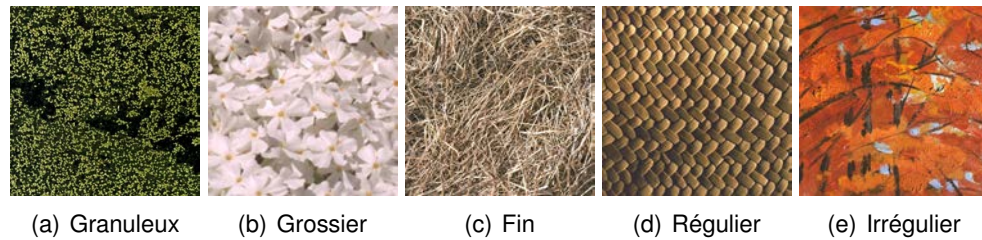


FIGURE 1.4 – Différents types de textures extraites de la base de texture VISTEX [VisTex 1995].

La texture est décrite dans la littérature par deux caractéristiques à savoir les contours (information mono-dimensionnelle) qui séparent les régions homogènes et la surface (information bidimensionnelle) qui définit ces régions homogènes. Cependant, cette description peut être erronée dès qu'un changement de résolution de l'image survient, ce qui entraîne la modification de l'échelle d'observation.

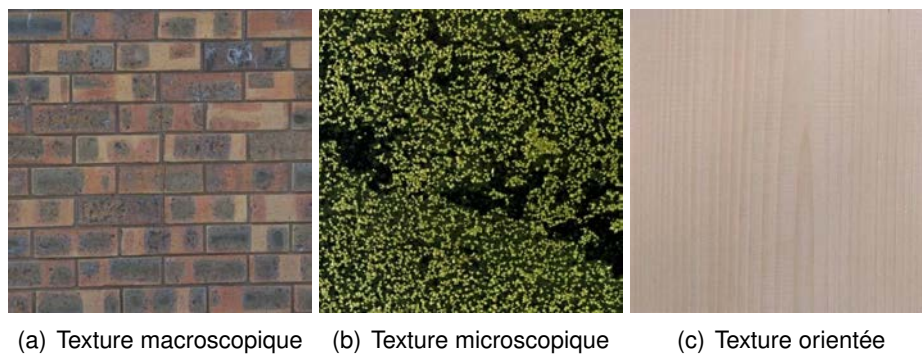


FIGURE 1.5 – Différentes catégories de texture.

Dans la littérature, un certain nombre de travaux [Haralick 1979, Cocquerez 1995] regroupent les approches de description de texture en deux principales familles. La première catégorie repose sur une description macroscopique de la texture (cf. figure 1.5(a)) basée sur la définition d'éléments primitifs répétitifs. Plus précisément, cette famille considère la texture comme étant un agencement spatial plus ou moins régulier d'éléments de base ou de primitives. Une primitive est une agglomération connexe de pixels ayant des propriétés similaires [Herlidou 1999]. Les partisans d'une telle description de texture [Tamura 1978, Haralick 1979, Julesz 1981, Herlidou 1999] s'orientent généralement vers des approches spectrales ou structurales qui sont des approches déterministes pour la caractériser. La deuxième famille d'approches est basée sur une étude microscopique de la texture (cf. figure 1.5(b)). Ces approches s'intéressent aux surfaces microscopiques isolables disposées généralement de manière anarchique et qui peuvent se répéter de manière discernable comme dans le cas du sable, gravier, herbe, etc. Les partisans [Sklansky 1978] d'une telle caractérisation de la texture s'orientent généralement vers des approches probabilistes. Dans la plupart des cas, la texture peut être ni totalement structurale ni totalement aléatoire [Mavromatis 2001, Hanbury 2002]. Ce type de texture se caractérise essentiellement par la répétition d'un motif de base selon certaines directions dominantes, sans qu'il y ait vraiment de répartition régulière ou complètement aléatoire. Cette texture est appelée

texture orientée. C'est le cas de l'image 1.5(c) issue de la base de texture de VISTEX.

Dans les images naturelles, la texture se présente sous la forme d'une mosaïque de différents types de textures. Par conséquent, la caractérisation d'un nombre infini de textures dans une image naturelle s'avère une tâche plus compliquée et moins précise, avec les outils usuels issus de la vision par ordinateur comparé à la capacité de l'œil humain. TUCERYAN [Tuceryan 1998] propose un ensemble de propriétés texturales intuitives exploitées dans l'analyse d'images en niveaux de gris. Les outils exploitant ces propriétés permettent d'étudier les impressions visuelles d'un observateur humain, ce qui permet de comprendre et de reproduire au mieux la caractérisation des différentes textures. Ces propriétés sont les suivantes :

- la texture peut être perçue à différents niveaux d'échelle<sup>1</sup> et à différents niveaux de résolution spatiale<sup>2</sup> ;
- la texture est un arrangement spatial des niveaux de gris des pixels ;
- l'étude de la texture demande la définition d'un voisinage spatial ayant une taille qui dépend du type de la texture ou de la surface de son motif de base ;
- si le nombre de motifs de base est suffisamment important dans une région, elle peut être décrite comme texturée.

La texture dans l'image peut être extraite en utilisant des outils mathématiques permettant d'obtenir des attributs de texture. Les paramètres calculés permettent donc de caractériser et d'analyser les images en question afin de les segmenter. Dans la littérature, il existe quatre classes connues d'attributs de texture [Tuceryan 1998, Porebski 2009] : les attributs géométriques, les attributs basés sur la modélisation spatiale des textures, les attributs spatio-fréquentiels et les attributs statistiques. Notons qu'un attribut peut satisfaire plusieurs propriétés. Dans cette thèse, nous constatons que nous pouvons considérer une cinquième catégorie d'attributs de texture calculés à l'aide des opérateurs de la morphologie mathématique (cf. figure 1.6). Nous détaillons dans la suite de ce chapitre, chaque catégorie d'attributs.

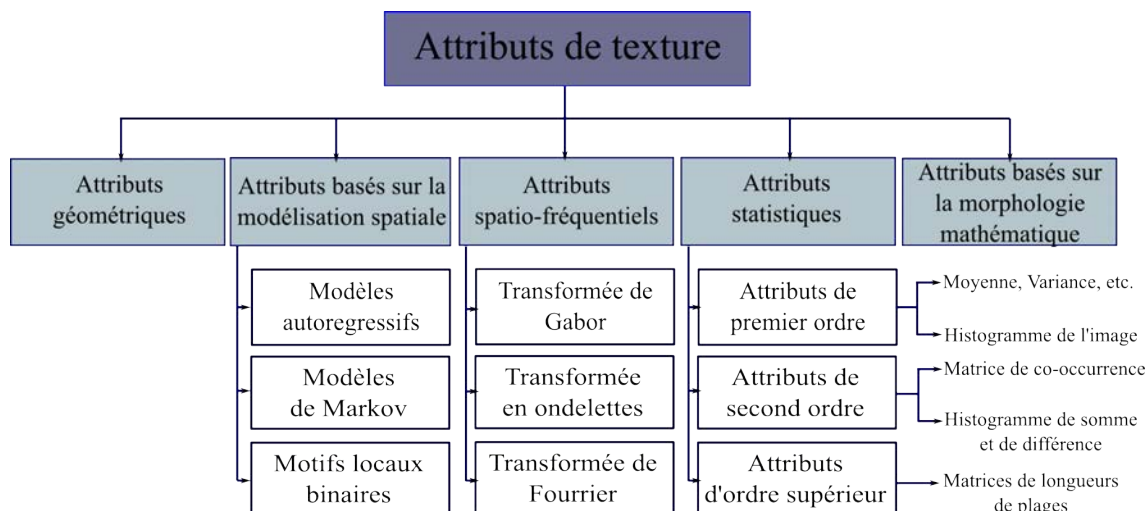


FIGURE 1.6 – Les différentes classes d'attributs de texture de la littérature.

1. L'échelle correspond ici au champ de vision : c'est la taille en millimètre des côtés de la surface observée

2. La résolution spatiale correspond au nombre de pixels par unité de longueur

### 1.2.2.1/ ATTRIBUTS GÉOMÉTRIQUES

La famille d'attributs géométriques est généralement adaptée à la perception macroscopique de la texture. Le calcul de ce type d'attributs commence par chercher des motifs de base (primitifs) de la texture et identifie ensuite des règles de placement de ces primitifs. En effet, les attributs géométriques exigent la régularité de la texture étudiée. Bien que des travaux de ZHENG [Zheng 2007] utilisent ce type d'attribut dans le but de caractériser la texture de la viande, il est à noter que cette famille d'attributs est très peu utilisée vu qu'elle ne prend pas en compte les irrégularités de la texture fortement présentes dans les images naturelles. Par conséquent, nous privilégions les attributs qui sont indépendants des conditions a priori faites sur la géométrie de la texture étudiée.

### 1.2.2.2/ ATTRIBUTS BASÉS SUR LA MODÉLISATION SPATIALE

Cette catégorie d'attributs dépend essentiellement de la répartition spatiale des couleurs dans l'image. Par conséquent, le calcul d'attributs est effectué en modélisant cette répartition spatiale. Parmi les modèles les plus répandus, nous pouvons citer les trois modèles suivants [Porebski 2009] :

#### a) Les modèles autorégressifs multispectraux

Les modèles autorégressifs multispectraux sont issus de l'extension des modèles autorégressifs en niveaux de gris. En effet, en un pixel  $p$ , les valeurs des composantes couleur sont obtenues en utilisant des combinaisons linéaires des niveaux des composantes couleur des pixels voisins auxquels un bruit blanc est ajouté. Les attributs de texture sont donc obtenus par estimation des paramètres de la combinaison linéaire. Le modèle Multispectral Simultaneous Auto-Regressive (MSAR) est le modèle le plus répandu pour des applications de segmentation de texture dans une image couleur [Hernandez 2004, Bennett 1998]. Le modèle MSAR est également utilisé pour la classification des bases de texture couleur [Hernandez 2005, Khotanzad 2006]. D'autres auteurs ont proposé l'extension des modèles autoregressifs, en séparant la luminance de la chrominance afin de caractériser les textures couleur de la base VisTex [Qazi 2009].

#### b) Les modèles de MARKOV

Les modèles de MARKOV constituent un cas particulier des modèles autoregressifs. L'intérêt de ces modèles (par rapport aux modèles autoregressifs) réside dans leur stratégie de modélisation des interactions spatiales entre un pixel avec ses voisins ce qui nécessite la définition d'un système de voisinage. En effet, pour définir les champs de Markov, l'image est modélisée en un ensemble fini  $S$  de sites  $s_i$  correspondant aux pixels. À chaque site, un descripteur est associé, ce qui permet de décrire l'état du site. Le descripteur peut être le niveau de gris ou une étiquette, mais il peut être une information plus complexe. Par exemple, pour les images couleur, le descripteur peut être une fonction paramétrique notée *fonction potentiel* pour laquelle les valeurs pour une clique<sup>3</sup> dépendent des relations entre les niveaux au sein d'une même composante couleur (relations *intra-composante*) et des relations entre les niveaux de deux composantes couleur différentes (relations *inter-composantes*). La définition d'interaction spatiale locale nécessite la structuration des relations spatiales entre ces

3. Un sous ensemble de  $S$ , relative au système de voisinage considéré et caractérisée par un ordre qui est son cardinal (le nombre d'éléments la constituant).

sites. Un système de cliques est déduit après avoir muni l'ensemble de sites  $S$  d'un système de voisinage  $\mathcal{V}$  défini par les équations suivantes [Tupin 2008] :

$$\mathcal{V}_s = \{t\} \text{ tels que } \begin{cases} s \notin \mathcal{V}_s \\ t \in \mathcal{V}_s \Rightarrow s \in \mathcal{V}_t \end{cases}$$

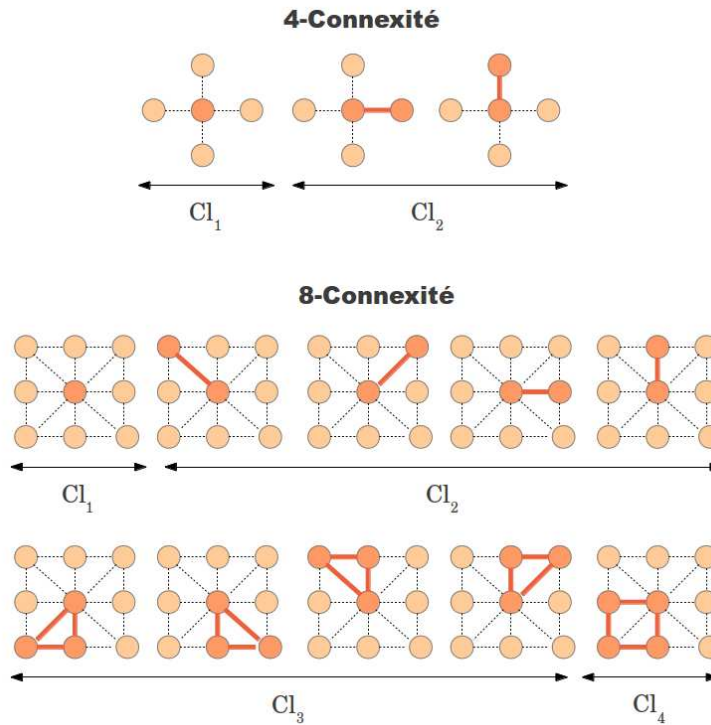


FIGURE 1.7 – Cliques.

La définition du voisinage permet de définir la notion de clique. En effet, une clique peut être soit un singleton de  $S$ , soit un ensemble de sites tous voisins les uns des autres. Dans la littérature, les voisinages les plus répandus dans le cas des champs de Markov sont le **4-voisinage** et le **8-voisinage**. Pour chaque voisinage, un système de cliques peut être déduit comme illustré dans la figure 1.7. Les modèles de MARKOV ont été largement utilisés par plusieurs auteurs à savoir KHOTANZAD [Khotanzad 2006] qui a estimé les paramètres de la *fonction potentiel* en utilisant les moindres carrées afin de caractériser et classer les textures couleur de la base VisTex, HERNANDEZ [Hernandez 2005] qui a mis en œuvre l'apport de la couleur lors de l'utilisation des modèles de MARKOV pour en calculer des attributs de texture au lieu d'utiliser seulement l'information de luminance et CHINDARO [Chindaro 2006] qui a utilisé des champs de Markov multispectral en combinant plusieurs espaces couleur.

### c) Les motifs locaux binaires

Les motifs locaux binaires (en anglais Local Binary Pattern : LBP) ont été initialement introduits par HARWOOD et al. [Harwood 1993] mais utilisés par OJALA et al. [Ojala 1996] pour analyser la texture. Cette technique consiste à indexer le motif local dans le voisinage d'un pixel  $p$  d'une image  $I$  [Porebski 2009, Pietikäinen 2011]. Plus précisément, les valeurs d'indexation sont obtenues en comparant les niveaux des composantes couleur du pixel courant  $p$  aux niveaux des composantes couleur de

ses voisins. Le résultat de la comparaison à chaque pixel voisin doit être codé en utilisant un masque connu par le masque de poids. Dans ce qui suit, nous détaillons la technique de calcul des valeurs de LBP en un pixel.

Soit  $LBP(p)^{C_k C_{k'}}$  le résultat de l'attribut LBP en un pixel donné  $p$  qui traduit la relation spatiale entre le pixel central (en prenant la valeur du niveau de gris de la composante  $C_k$ ) et ses pixels voisins (en prenant la valeur de leur niveau de gris de la composante couleur  $C_{k'}$ ). Pour obtenir la valeur de  $LBP(p)^{C_k C_{k'}}$ , la première étape consiste à effectuer un seuillage des niveaux de gris des pixels voisins (de la composante couleur  $C_{k'}$ ) dans un voisinage  $\mathcal{V}_p$ , en utilisant comme seuil la valeur  $C_k(p)$  (cf. équation 1.1).

$$\forall t \in \mathcal{V}_p \begin{cases} \text{Si } C_{k'}(t) \leq C_k(p) & \text{alors } LBP(p)^{C_k C_{k'}} = 0 \\ \text{Si } C_{k'}(t) \geq C_k(p) & \text{alors } LBP(p)^{C_k C_{k'}} = 1 \end{cases} \quad (1.1)$$

où  $k, k' \in \{1, 2, 3\}$  pour un espace couleur donné  $(C_1, C_2, C_3)$ .

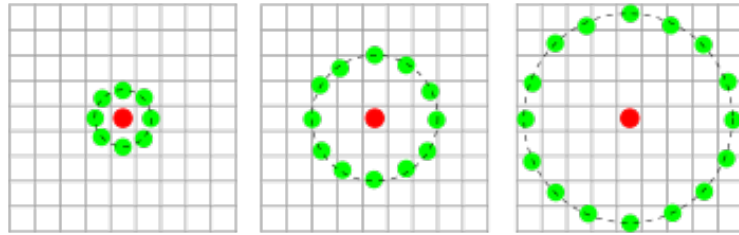


FIGURE 1.8 – Exemples de voisinages utilisés pour le calcul des LBP : à gauche  $N = 8$ , au centre  $N = 12$  et à droite  $N = 16$ .

La deuxième étape consiste à appliquer le masque de poids aux résultats de seuillage. En effet, il suffit de multiplier les valeurs des pixels voisins, obtenues après le seuillage, par les valeurs du masque. Le masque utilisé est constitué de valeurs prises  $2^0$  et  $2^{N-1}$ , où  $N$  est le nombre des voisins considérés pour le pixel  $p$ . Notons que le voisinage du pixel considéré peut être distribué de plusieurs manières comme illustré sur la figure 1.8. Pour une meilleure illustration des étapes de calcul du LBP (seuillage et masquage), nous présentons un exemple sur la figure 1.9. Étant donné une image de texture couleur dans un espace couleur  $(C_1, C_2, C_3)$ , nous obtenons, pour chaque pixel donné  $p$ , neuf coefficients de  $LBP$  dont trois caractérisent les relations *intra-composantes* ( $LBP(p)^{C_k C_k} \forall k \in \{1, 2, 3\}$ ) et six caractérisent les relations *inter-composantes* ( $LBP(p)^{C_k C_{k'}} \forall k \neq k' \in \{1, 2, 3\}$ ). Les coefficients de  $LBP$  calculés en chaque pixel sont, tout d'abord, rassemblés dans 9 histogrammes. Ces derniers sont, ensuite, concaténés pour obtenir un histogramme final, noté  $LBP(I)$ , qui caractérise la texture complète [Pietikäinen 2002].



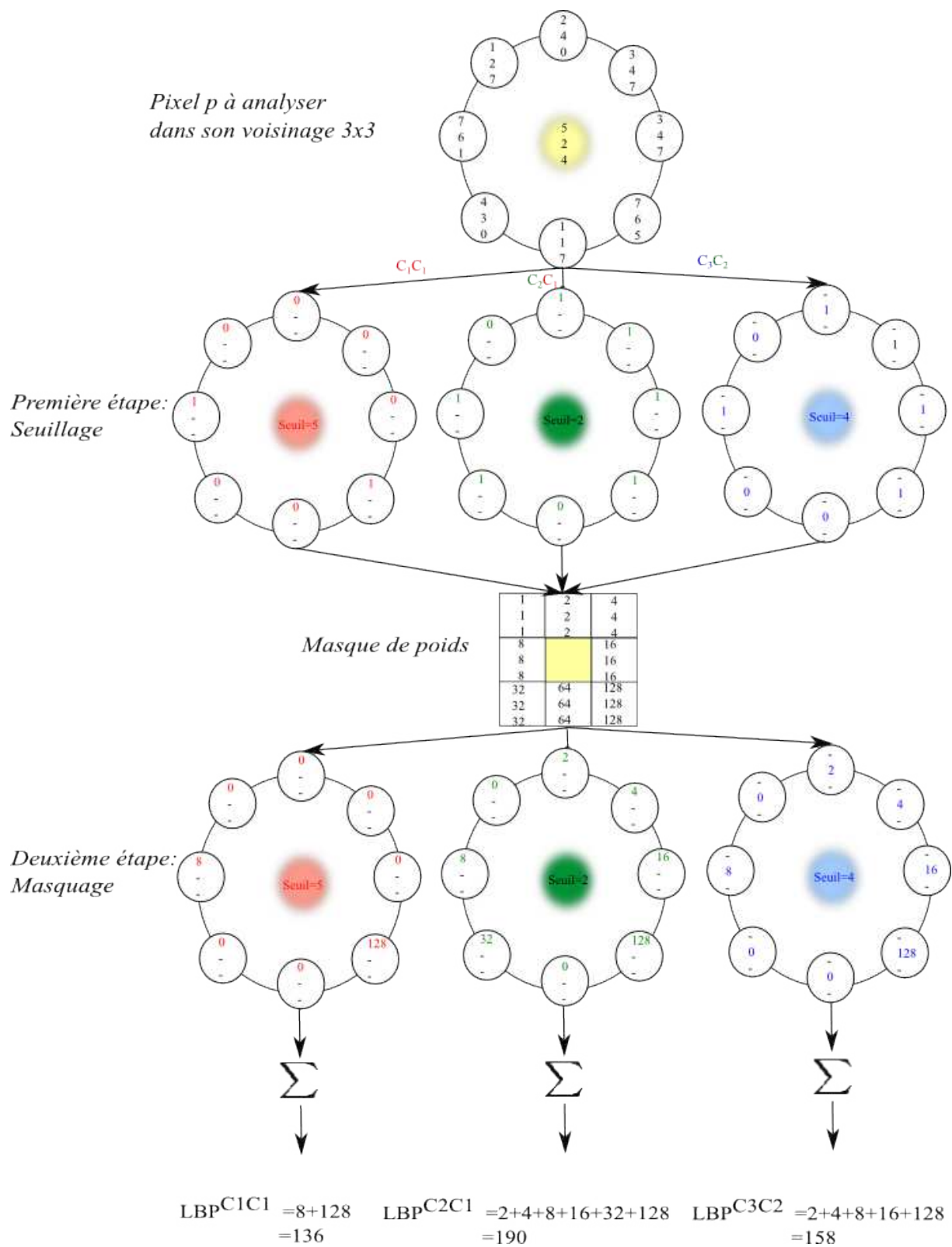


FIGURE 1.9 – Illustration du principe de calcul du LBP.

Tout comme les modèles autorégressifs et les modèles de MARKOV, les LBP font partie des attributs de texture utilisant la modélisation spatiale. Il existe d'autres attributs caractérisant la texture couleur que nous présentons dans la suite de cette section.

### 1.2.2.3/ ATTRIBUTS SPATIO-FRÉQUENTIELS

Les attributs spatio-fréquentiels se divisent en trois catégories : ceux qui se calculent dans le domaine spatial, dans le domaine fréquentiel ou bien dans le domaine spatio-fréquentiel. Nous détaillons dans ce qui suit chaque catégorie :

#### a) Dans le domaine spatial

Le calcul d'attributs de texture dans le domaine spatial consiste à caractériser la texture par le nombre de transitions par unité de surface. Par conséquent, contrairement aux textures fines, les textures grossières tendent à avoir une faible densité de transition [Tuceryan 1998]. Par exemple, les transitions couleur peuvent être détectées par la méthode de DI-ZENZO [DiZenzo 1986].

#### b) Dans le domaine fréquentiel

Le calcul d'attributs de texture dans le domaine fréquentiel consiste à effectuer une transformation du domaine spatial au domaine fréquentiel. Les transformations les plus répandues sont effectuées par une transformée de Fourier et une transformée en cosinus. Celles-ci sont basées sur l'utilisation d'un ensemble de filtres sensibles chacun à une fréquence particulière. Notons que dans le cas des images couleur, le filtrage est effectué sur chaque composante couleur. À partir de l'image résultante (image filtrée), les attributs de texture sont calculés en utilisant des outils statistiques à savoir l'énergie et la variance. Il est à noter que la transformée de fourrier et la transformée en cosinus sont souvent mieux adaptées pour des images ayant des zones homogènes (texture grossière). Par conséquent, l'intégralité de l'information est représentée par les basses fréquences tandis que les hautes fréquences sont réservées à des changements locaux.

#### c) Dans le domaine spatio-fréquentiel

Un troisième type d'attributs de texture associant les deux domaines spatial et fréquentiel existe. Parmi les transformations issues de cette catégorie et communément utilisées, nous pouvons citer la transformée de Gabor [Palm 2002] et la transformée d'ondelettes [Carré 2001].

Maintenant que nous avons rappelé brièvement les définitions des attributs de texture géométriques, basés sur la modélisation spatiale et spatio-fréquentielle, nous introduisons dans ce qui suit, une quatrième famille d'attributs de texture qui sont les attributs statistiques.

### 1.2.2.4/ ATTRIBUTS STATISTIQUES

Les attributs statistiques traduisent la relation (en termes de variation de couleur) entre un pixel et ses voisins en tenant compte d'un système de voisinage défini a priori. Les attributs statistiques peuvent être divisés en trois ensembles à savoir les attributs statistiques de premier ordre tel que seulement le pixel en question est pris en compte (les histogrammes), les attributs statistiques de deuxième ordre (les matrices de co-occurrences et les histogrammes de sommes et de différences) et les attributs

statistiques d'ordre supérieure (les matrices de longueurs de plages). Notons que les attributs statistiques sont très utilisés grâce à la pertinence de leurs outils pour caractériser même les textures fines qui ne présentent pas de régularités apparentes. Nous rappelons, dans la suite, les attributs statistiques de texture, de chaque ensemble, les plus communément utilisés.

#### a) Attributs statistiques de premier ordre

Les attributs statistiques de premier ordre renferment deux ensembles. Le premier ensemble regroupe les statistiques calculées directement sur l'image couleur à savoir la moyenne des composantes couleur, la variance qui mesure la dispersion des composantes couleur autour de leur moyenne, la dissymétrie qui mesure l'asymétrie des composantes couleur à leur moyenne et l'aplatissement qui mesure le degré de convexité ou de concavité des composantes couleur par rapport à leur moyenne. Ces attributs statistiques que nous venons de citer sont peu performants vu qu'ils n'exploitent que l'information couleur de l'image sans tenir compte des interactions spatiales entre les pixels.

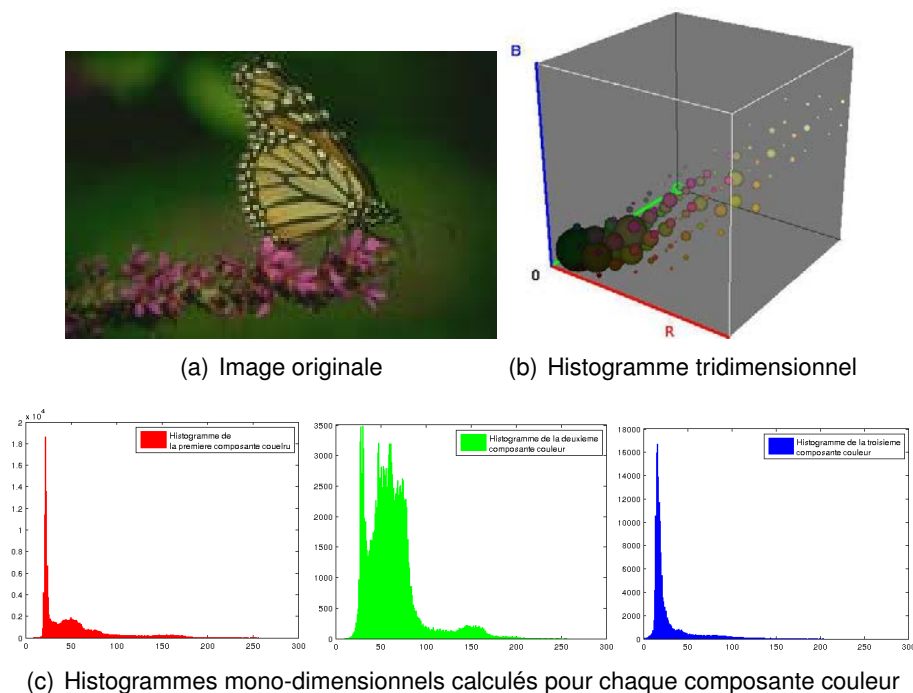


FIGURE 1.10 – Illustration des différents histogrammes mono-dimensionnels et tridimensionnel d'une image couleur dans l'espace couleur ( $R, G, B$ ) (de haut en bas et de gauche à droite : image originale couleur, histogramme tridimensionnel obtenue par le logiciel **Color Inspector 3D** et histogrammes mono-dimensionnels de la première, deuxième et troisième composante couleur).

Le deuxième ensemble d'attributs statistiques de texture de premier ordre sont les histogrammes d'image. Ces derniers sont des outils permettant de décrire la distribution des niveaux de gris des composantes couleur de l'image. Il existe deux formes d'histogramme à savoir l'histogramme mono-dimensionnel calculé pour chaque composante couleur et l'histogramme tridimensionnel représenté généralement sous forme de cellules où chacune d'elles contient l'occurrence de



pixels ayant une même couleur. À titre indicatif, si les composantes couleur sont codées sur 8 bits, l'histogramme de l'image contiendra  $256^3$  cellules. La figure 1.10 illustre l'exemple des différents histogrammes mono-dimensionnels (cf. figure 1.10(c)) et tridimensionnel (cf. figure 1.10(b)) d'une image couleur dans l'espace couleur  $(R, G, B)$ .

Dans la pratique et dans la plupart des applications de segmentation d'images, les histogrammes ne sont pas directement utilisés. En effet, des nouveaux paramètres statistiques sont calculés à partir de ces histogrammes à savoir la médiane qui permet de découper l'histogramme en deux parties d'égales populations, le mode qui correspond au maximum de l'histogramme et l'intervalle inter-quartile qui caractérise la dispersion du niveau de gris autour de la médiane. Bien que ces attributs sont faciles à déterminer et à exploiter, leur inconvénient majeur réside dans l'exploitation unique de l'information couleur sans tenir compte des interactions spatiales entre pixels de l'image.

#### b) Attributs statistiques de second ordre

Les deux attributs statistiques de second ordre les plus répandus sont les matrices de co-occurrences et les histogrammes de somme et de différence. Les matrices de co-occurrences, proposées par HARALICK en 1973 [Haralick 1973], constituent un outil performant pour la description des interactions spatiales entre les pixels. En effet, cet outil permet d'étudier la distribution de l'information couleur, notamment dans le cas des matrices de co-occurrences chromatiques [Palm 2004], tout en tenant compte de la relation spatiale entre les pixels. Nous nous intéressons dans ce qui suit aux matrices de co-occurrences chromatiques qui ont prouvé leur pertinence et efficacité pour la segmentation d'images, en comparaison avec les matrices monochromatiques.

Les matrices de co-occurrences d'une image couleur  $I$ , dans un espace couleur  $(C_1, C_2, C_3)$ , sont déterminées en calculant l'occurrence d'un couple de niveaux de gris dont la relation spatiale est définie par un vecteur de déplacement  $\vec{R}$ . Cette relation spatiale peut être aussi définie par un angle  $\theta$  et une mesure  $\|\vec{R}\|$  donnant respectivement la direction et le module du vecteur  $\vec{R}$  (la distance spatiale en nombre de pixel entre le pixel à analyser et ses voisins) comme illustré sur la figure 1.11.

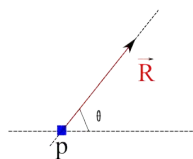


FIGURE 1.11 – Un vecteur de translation  $\vec{R}$  en fonction des paramètres  $\theta$  et  $\|\vec{R}\|$ .

Soit  $M^{C_k C_{k'}}[I](\vec{R})$ ,  $\forall k, k' \in \{1, 2, 3\}$ , la matrice de co-occurrences mesurant l'interaction spatiale entre deux pixels de l'image  $I$ . Pour une paire de niveaux de gris  $(m, n)$  correspondant à la cellule d'indices  $(m, n)$  dans la matrice de co-occurrences, la valeur de cette cellule est égale à l'occurrence des couples de niveaux de gris formés par : le niveau de gris de la composante couleur  $C_k$  de  $I$  en un pixel  $p$  telle que  $I^{C_k}(p) = m$  et le niveau de gris d'un pixel obtenu après la translation de  $p$  par le vecteur  $\vec{R}$  telle que  $I^{C_{k'}}(p) = n$  (cf. équation 1.2).

$$M^{C_k C_{k'}}[I](\vec{R})(m, n) = \text{card}\{(I^{C_k}(p), I^{C_{k'}}(p + \vec{R})) = (m, n)\} \quad (1.2)$$

Dans la littérature, l'angle  $\theta$  prend souvent ses valeurs dans l'ensemble  $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ . La figure 1.12 illustre des exemples de translation par différents vecteurs  $\vec{R}$ .

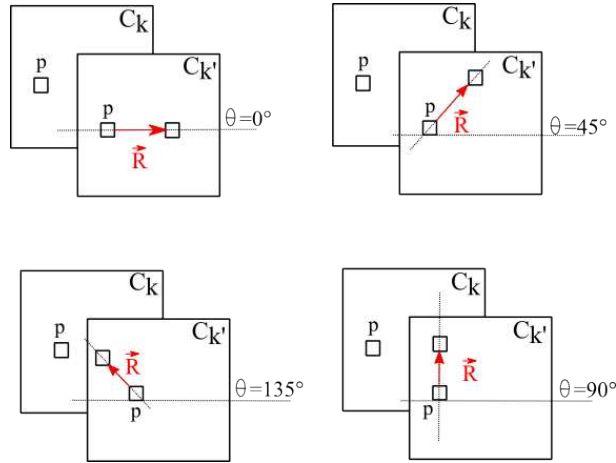


FIGURE 1.12 – Exemple de différents vecteurs de translation  $\vec{R}$ .

Nous présentons maintenant, un deuxième attribut statistique de second ordre, moins répandu que les matrices de co-occurrences : les histogrammes de somme et de différence, proposés par UNSER [Unser 1986]. Tout comme les matrices de co-occurrences, les histogrammes de somme et de différence ont été initialement utilisés pour les images en niveaux de gris, puis étendus dans le cas des images couleur. Cet attribut présente des performances d'analyse du contenu de l'image quasi-similaires en comparaison avec les matrices de co-occurrences, avec l'avantage d'occuper un espace mémoire et un temps de calcul plus réduits.

Le principe de calcul des histogrammes de somme et de différence ressemble à celui des matrices de co-occurrences. Soit  $I$  une image couleur codée dans un espace couleur  $(C_1, C_2, C_3)$ . L'histogramme de somme (respectivement de différence) consiste à calculer l'occurrence de la somme (respectivement de la différence) du niveau de gris dans la composante couleur  $C_k$  d'un pixel  $p$  et le niveau de gris dans la composante couleur  $C_{k'}$  du pixel après une translation  $\vec{R}$ . Notons  $H_S^{C_k C_{k'}}[I](\vec{d})$  l'histogramme de somme et  $H_D^{C_k C_{k'}}[I](\vec{d})$  l'histogramme de différence de l'image  $I$  entre les composantes couleur  $C_k$  et  $C_{k'}$ ,  $\forall k, k' \in \{1, 2, 3\}$ . Pour une cellule d'indice  $m \in [0, 2 \times (N - 1)]$  de l'histogramme de somme où  $N$  est le niveau de quantification, sa valeur est donnée par l'équation 1.3. Pour une cellule d'indice  $n \in -(N - 1), (N - 1)[$  de l'histogramme de différence, sa valeur est donnée par l'équation 1.4.

$$H_S^{C_k C_{k'}}[I](\vec{d})(m) = \text{card}\{I^{C_k}(p) + I^{C_{k'}}(p + \vec{d}) = m\} \quad (1.3)$$

$$H_D^{C_k C_{k'}}[I](\vec{d})(n) = \text{card}\{I^{C_k}(p) - I^{C_{k'}}(p + \vec{d}) = n\} \quad (1.4)$$

À partir de ces deux outils statistiques de second ordre, nous pouvons obtenir pour une image donnée, une collection de matrices de co-occurrences ou une collection

d'histogrammes de somme et de différence pour différents vecteurs de translation. Cependant, dans la pratique, ces collections ne sont pas directement exploitées pour caractériser la texture d'une image. Par conséquent, HARALICK [Haralick 1973] propose quatorze nouveaux paramètres calculés à partir des matrices de co-occurrences tandis que UNSER [Unser 1986] propose seize nouveaux paramètres calculés à partir des histogrammes de somme et de différence.

Bien que les histogrammes de somme et de différence soient moins coûteux en temps de calcul et en espace mémoire, les matrices de co-occurrences restent l'attribut statistique de second ordre le plus utilisé dans la littérature [Porebski 2007, Tesar 2007, Meurie 2010a, Niea 2011, Dash 2012]. Pour cette raison, nous avons fait le choix d'utiliser les matrices de co-occurrences pour caractériser l'information de texture utilisée ensuite dans la stratégie de combinaison proposée. Les paramètres d'HARALICK issus des matrices de co-occurrences et qui seront utilisés dans nos expérimentations, seront détaillés dans le chapitre 3. Le lecteur intéressé par les paramètres issus des histogrammes de somme et de différence pourra se référer à [Porebski 2009] pour plus de détails.

### c) Attributs statistiques d'ordre supérieur

L'attribut statistique d'ordre supérieur le plus connu la matrice de longueurs de plages. Cet attribut est déterminé par le calcul de l'occurrence des niveaux de gris des composantes couleur d'un segment de pixels (au lieu d'un couple de pixels dans le cas des matrices de co-occurrences). Les pixels appartenant au segment considéré doivent être adjacents, consécutifs et avoir le même niveau de gris dans une composante couleur  $C_k$  selon une direction donnée [Zheng 2007]. Soit  $I$  une image couleur dans l'espace couleur  $(C_1, C_2, C_3)$ . Une matrice de longueurs de plages  $\mathcal{P}^{C_k}[I](\theta, (m, L))$  est calculée en fonction de la direction  $\theta$  (souvent parmi  $\{0^0, 45^0, 90^0$  et  $135^0\}$ ), de la composante couleur  $C_k$  ( $\forall k \in \{1, 2, 3\}$ ), de  $m$  le niveau de gris dans une composante couleur  $C_k$  des différents pixels formant le segment considéré et de  $L$  la longueur en nombre de pixel de ce segment. Ainsi, une cellule de cette matrice calcule l'occurrence de la suite de pixels adjacents dans une direction  $\theta$  et ayant le niveau de gris  $m$  d'une composante couleur  $C_k$ . Tout comme les matrices de co-occurrences et les histogrammes de somme et de différence, les matrices de longueurs de plages ne sont pas directement exploitées. Pour cette raison, GALLOWAY [Galloway 1975] propose cinq paramètres extraits des matrices de longueurs de plages (détaillés dans les travaux de POREBSKI [Porebski 2009]).

La caractérisation de la texture dans une image peut être effectuée également grâce à d'autres outils issus de la morphologie mathématique. Leur mode de calcul étant fondé sur la théorie des ensembles, nous avons choisi de regrouper ces outils dans une cinquième catégorie : les attributs de texture issus de la morphologie mathématique.

La morphologie mathématique est une nouvelle théorie d'analyse d'image inventée dans les laboratoires de l'école des mines de Paris en 1964. Les attributs issus de la morphologie mathématique s'appuient sur la théorie des ensembles, les treillis, la topologie des ensembles fermés et les probabilités. Ces attributs peuvent ainsi s'appliquer à des structures algébriques variées (des fonctions, des ensembles flous ou des propositions logiques). Ils peuvent avoir, dans certains cas, un caractère aléatoire comme dans le cas d'analyse de texture [Angelini 2007]. La plupart des attributs définis dans la morphologie mathématique sont le résultat de combinaisons d'opérateurs de

base de cette théorie à savoir l'érosion, la dilatation, l'ouverture et la fermeture. Parmi les opérateurs de la morphologie mathématique permettant l'étude de la texture dans une image, nous pouvons citer le filtre alterné séquentiel (ASF) et le gradient morphologique. Nous retenons ces deux derniers pour nos expérimentations réalisées dans cette thèse.

Dans cette section, nous avons rappelé les définitions de la couleur et de la texture et présenté brièvement différents types d'attributs permettant de caractériser la texture (fine, grossière, régulière, irrégulière, orientée, etc.) dans une image. La question qui se pose est : Est ce que la couleur seule ou la texture seule peut satisfaire les besoins de l'utilisateur pour décrire/identifier le contenu d'une image ?

Depuis longtemps, la couleur et la texture sont deux informations exploitées de manière disjointe dans de nombreux travaux sur la segmentation d'images. Les travaux récents prouvent que la combinaison des informations de couleur et de texture pour la caractérisation et l'identification des objets dans une image conduit à des résultats beaucoup plus intéressants et précis. Nous présentons donc, dans la section suivante, les principales méthodes de segmentation d'images issues de la littérature qui utilisent conjointement la couleur et la texture. Ces méthodes seront utilisées dans l'étude comparative menée pour évaluer notre méthode de combinaison couleur/texture appliquée à la segmentation d'images.

### 1.3/ SEGMENTATION D'IMAGES PAR UTILISATION CONJOINTE DE LA COULEUR ET DE LA TEXTURE

Les méthodes de segmentation d'images utilisant conjointement la couleur et la texture permettent de partitionner une image en étant plus proche de la perception humaine que celles qui utilisant la couleur ou la texture seule.

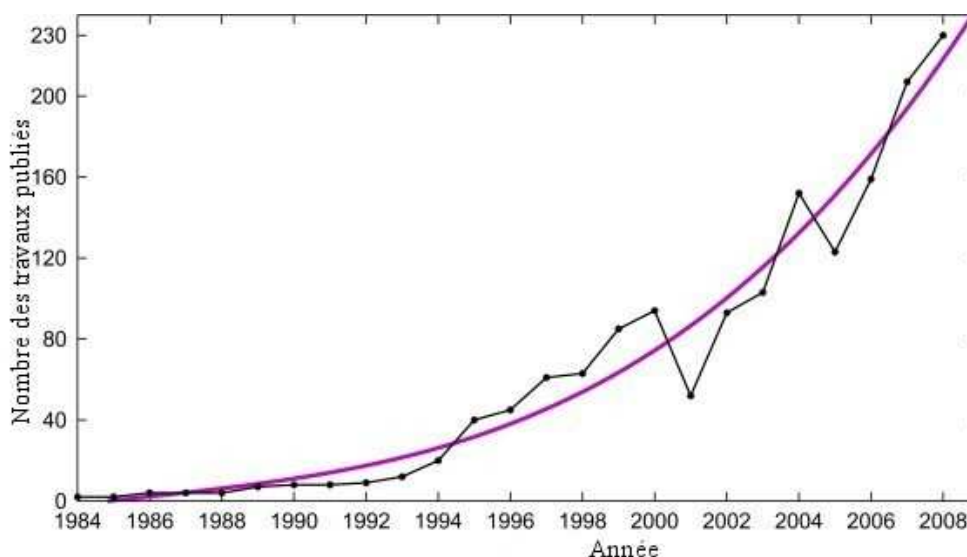


FIGURE 1.13 – Taux (en noir) et direction (en violet) de croissance des travaux publiés sur les méthodes de segmentation couleur/texture entre les années 1984 et 2008 [Ilea 2011].

D'une manière générale, la combinaison des informations de couleur et de texture est intéressante dans la mesure où elle permet de concourir à améliorer les résultats de segmentation en comparaison à l'utilisation seule de l'une des deux sources. Nous suivons dans le reste de ce chapitre le plan de présentation tel qu'on le trouve dans les travaux de ILEA [Ilea 2011] qui réalise une étude bibliographique détaillée sur les méthodes de segmentation intégrant la couleur et la texture entre les années 1984 et 2010 (cf. figure 1.13).

Dans cette section, nous présentons une synthèse des principales méthodes de segmentation couleur/texture. Dans la littérature, nous pouvons distinguer trois catégories d'utilisation conjointe de la couleur et de la texture et que nous avons regroupées sur la figure 1.14 : 1/ l'intégration implicite des attributs couleur et des attributs texture ; 2/ les méthodes qui extraient l'information couleur et l'information texture en cascade ; 3/ les approches qui extraient les attributs couleur et les attributs texture séparément et les combinent lors du processus de segmentation. Dans ce qui suit, nous présentons un certain nombre de méthodes de de segmentation, issues de la littérature, se positionnant dans chacune de ces catégories.

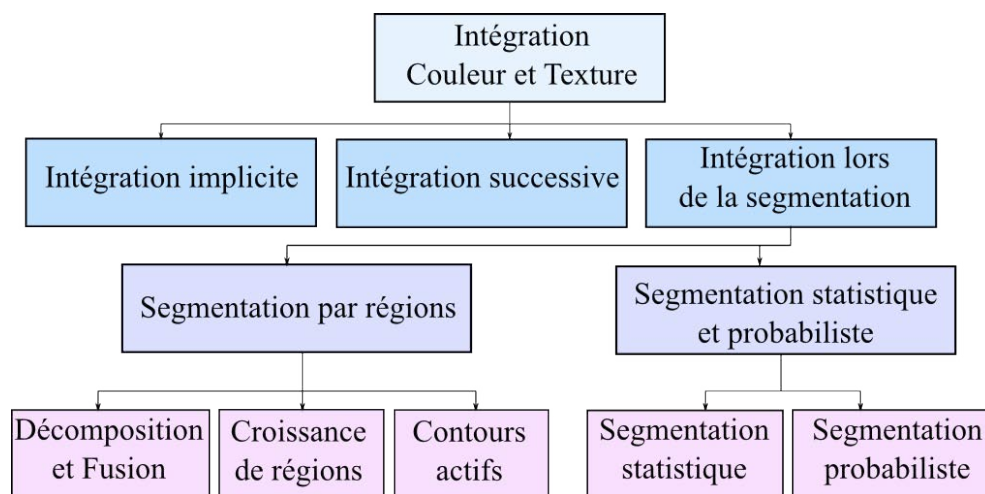


FIGURE 1.14 – Les différents types des méthodes de segmentation utilisant conjointement la couleur et la texture.

### 1.3.1/ INTÉGRATION IMPLICITE

Dans le cadre d'une intégration implicite, l'information texture et l'information couleur sont mutuellement dépendantes. Par conséquent, leur extraction se fait suivant trois cas différents : 1/ à partir d'une seule couche couleur ; 2/ à partir d'une paire de composantes couleur corrélées ; ou bien 3/ à partir d'une combinaison de composantes couleur. Nous rappelons donc ci-après, le concept des méthodes de segmentation les plus connues, faisant partie de cette catégorie, à savoir les méthodes CTM (Compression-based Texture Merging), MeanShift et Ncuts (Normalized cuts).

### 1.3.1.1/ COMPRESSION-BASED TEXTURE MERGING (CTM)

La méthode de segmentation Compression-based Texture Merging (CTM), introduite dans les travaux de YANG et al. [Yang 2008], consiste à traiter le problème de segmentation d'images comme un problème de classification d'attributs couleur/texture (considérés comme des données mélangées multivariées). L'image initialement codée dans un espace couleur donné, doit être convertie dans l'espace couleur CIE Lab, représentant l'espace couleur le plus uniforme selon les auteurs. Ces derniers modélisent la distribution des attributs couleur/texture avec un mélange de distributions gaussiennes grâce à la représentation de l'image dans l'espace couleur CIE Lab. Cet espace, contrairement à l'espace couleur  $(R, G, B)$  où la relation entre les trois coordonnées est non linéaire, facilite la représentation de l'information texture sous la forme d'une gaussienne. Contrairement à la plupart des méthodes de segmentation, les auteurs supposent que le mélange des composantes est dégénéré<sup>4</sup> ou bien quasi-dégénéré. Cette hypothèse est particulièrement importante pour la segmentation d'images à mi-niveau. Les auteurs montrent donc qu'un mélange de distributions peut être efficacement classé grâce à un algorithme de classification ascendant dérivé à partir d'une approche de compression de données avec perte. En effet, pour chaque composante couleur de l'espace CIE Lab, les valeurs de l'intensité sont stockées dans une fenêtre de taille fixe ( $7 \times 7$  pixels correspond à la taille de la fenêtre retenue pour des résultats satisfaisants) centrée en chaque pixel de l'image sur laquelle est ensuite appliqué un filtre gaussien. Finalement, la dimension du vecteur de descripteurs est réduite à huit à l'aide d'une analyse en composantes principales (ACP). Une illustration de la construction du vecteur d'attributs est donnée par la figure 1.15. La méthode CTM a été appliquée et évaluée sur des images de la base de BERKELEY [Berkeley].

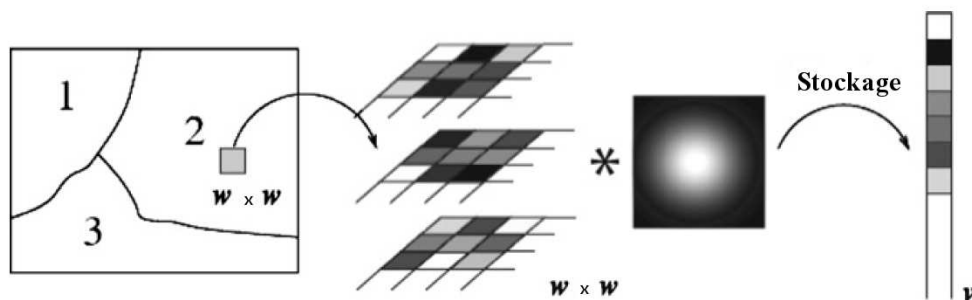


FIGURE 1.15 – Construction d'un vecteur d'attributs couleur/texture. Une fenêtre de taille  $w \times w$  de chaque composante couleur de l'espace CIE Lab est filtrée par un filtre gaussien. Les valeurs obtenues pour chaque composante couleur sont stockées dans un seul vecteur [Yang 2008].

### 1.3.1.2/ MEANSHIFT

L'algorithme MeanShift introduit par COMANICIU [Comaniciu 2002] est une méthode de segmentation assez répandue. Dans cet algorithme, le pixel est représenté par la concaténation des coordonnées spatiales et de l'information radiométrique qui sont sauvegardées dans un seul vecteur. Cette méthode consiste à appliquer deux étapes

4. La dégénérescence est introduite en utilisant une représentation commune d'attributs de différentes textures dans une image.

principales à savoir le filtrage MeanShift qui permet de donner une convergence à chaque pixel et la segmentation MeanShift qui consiste à grouper les pixels ayant une convergence inférieure à deux seuils : un dans le domaine spatial (spatial domain) et un autre dans le domaine radiométrique (range domain).

#### a) Filtrage MeanShift

L'étape de filtrage a été initialement introduite par FUKUNAGA en 1975 afin d'estimer le gradient de la densité d'un nuage de points. Cette technique a été exploitée par FUKUNAGA pour des problèmes de reconnaissance de formes [Fukunaga 1975] et a été retravaillée par COMANICIU pour être utilisée dans plusieurs autres travaux [Comaniciu 1999, Comaniciu 2001, Comaniciu 2002]. L'étape de filtrage MeanShift consiste à appliquer une procédure itérative de montée de gradient permettant l'estimation de modes principaux<sup>5</sup> d'une fonction de densité notée  $f(x)$  d'un nuage de points. Par conséquent, le filtrage MeanShift revient à chercher de manière itérative les solutions de l'équation  $\nabla f(x) = 0$  sans estimer la densité de  $f$ . Cette étape de filtrage étant décrite en détail dans [Comaniciu 2002], nous nous contentons de rappeler brièvement son principe. Le filtrage MeanShift est décrit par l'algorithme 1 dont la convergence a été vérifiée dans [Comaniciu 2002]. Les points  $y_j$  calculés à chaque itération sont appelés les points de trajectoire tel que le premier point noté  $y_1$  constitue le point initial d'estimation et le dernier point constitue le mode cherché.

---

##### Algorithme 1: Filtrage MeanShift

---

**Entrées** : Soient  $\{p_i\}_{i=1\dots n}$ ,  $n$  pixels dans l'espace de dimension  $d$  et  $\{z_i\}_{i=1\dots n}$  leurs modes associés

**Pour**  $i = 1$  à  $n$  **faire**

    Initialiser  $j = 1$ ,  $y_1 = p_i$

**Répéter**

        Calculer  $m(y_j)$  le vecteur MeanShift comme défini [Comaniciu 2002]

$y_{j+1} = y_j + m(y_j)$

$j = j + 1$

**Jusqu'à**  $y_{j-1} = y_j$ ;

$z_i = y_j$

**fin**

**Sorties** : Les modes  $z_i$

---

La procédure de filtrage nécessite une fonction noyau pour définir le gradient qui peut influencer les résultats de la convergence et la qualité d'estimation. Des travaux [Bugeau 2011] ont montré que le noyau d'EPANECHNIKOV converge plus rapidement qu'un noyau gaussien, tandis que ce dernier possède une qualité d'estimation meilleure.

#### b) Segmentation MeanShift

L'étape de segmentation MeanShift se servant de l'étape de filtrage précédente, consiste à regrouper tous les points convergeant vers le même mode. L'algorithme 2, présenté ci-après, décrit le principe de segmentation MeanShift. Une étape supplémentaire mais optionnelle peut être ajoutée afin de ne conserver que les régions dont la surface est supérieure à un minimum donné.

---

5. Un mode est un maximum local de la fonction de densité d'un nuage de points

---

**Algorithme 2:** Segmentation MeanShift

---

**Entrées :** Soient  $\{p_i\}_{i=1\dots n}$ ,  $n$  pixels dans l'espace de dimension  $d$  et  $\{z_i\}_{i=1\dots n}$  leurs modes associés.

Appliquer l'algorithme de filtrage MeanShift.

Regrouper tous les modes  $z_i$  séparés d'une distance inférieure à  $h$ , cela revient à effectuer :

**Pour**  $i, j = 1$  à  $n$  **faire**

**Si** Deux modes  $z_i$  et  $z_j$  vérifient  $\|z_i - z_j\| \leq h$  **alors**

        Mettre  $z_i$  et  $z_j$  dans le même groupe.

**fin**

**fin**

Regrouper tous les points  $p_i$  dont les modes appartiennent au même groupe.

---

### 1.3.1.3/ NORMALIZED CUTS (NCUTS)

La méthode de segmentation Normalized Cuts (NCuts) proposée dans les travaux de SHI et MALIK [Shi 2000], est basée sur la théorie spectrale de graphes. Contrairement aux approches basées sur l'analyse locale de l'image, elle permet l'intégration des informations couleur/texteure de l'image grâce à un processus de regroupement de graphe. Dans ce contexte, les auteurs traitent la segmentation d'images comme un problème de partitionnement de graphe tout en définissant un nouveau critère global Normalised Cuts qui donne son nom à la méthode proposée. Le critère défini mesure aussi bien la dissimilarité entre les différentes régions que la similarité totale dans ces régions. La méthode de NCuts a été étendue pour être appliquée à la segmentation des images multi-échelle [Cour 2005] après avoir été définie pour des images monochromatiques. L'extension de la méthode NCuts aux images multi-échelles a été effectuée en propageant le regroupement d'informations locales en multiples valeurs de connections spatiales. Pour plus de détails sur cette extension, le lecteur pourra se référer aux travaux de COUR et al. [Cour 2005].

---

**Algorithme 3:** Normalized Cuts

---

**Entrées :** Une image ou une séquence d'images

- 1 Établir un graphe  $G = (V, E)$ .
  - 2 Définir la matrice de similarité  $W$  (dans d'autres travaux on parle de matrice d'affinité)
  - 3 Déterminer les vecteurs propres correspondant aux plus petites valeurs propres de la matrice  $W$ .
  - 4 Partitionner le graphe en deux parties à l'aide du vecteur propre correspondant à la deuxième plus petite valeur propre.
  - 5 Décider si la partition courante doit être divisée une autre fois et répartir récursivement les parties segmentées.
- 

Considérons  $I$  l'image initiale et  $G = (V, E, W)$  le graphe construit associé.  $V$  est le graphe de nœuds formé par les pixels de l'image.  $E$  est le graphe de contours reliant les pixels séparés d'une distance inférieure à un seuil  $G_r$ .  $W$  définit la matrice d'affinité qui mesure la similarité entre les pixels tel que  $w(i, j)$  exprime la similarité entre les pixels  $p_i$  et  $p_j$  appartenant à la même région de l'image. La définition de la matrice d'affinité  $W$  ainsi que la valeur optimale du critère du Normalised Cuts sont présentés avec plus de



détails dans les travaux de [Shi 2000, Cour 2005]. Dans cette section, nous rappelons brièvement les principales étapes de la méthode de segmentation NCuts en présentant l'algorithme 3.

Il existe d'autres méthodes de segmentation opérant selon l'intégration implicite [Panjwani 1993, Paschos 1995, Shafarenko 1997, Hoang 2005, Wang 2006, Shi 2007]. Nous citons sans rentrer trop dans les détails les travaux de PANJWANI et HEALEY [Panjwani 1993] qui proposent une méthode basée sur les champs de Markov GMRF (Gaussian Markov Random Field). La méthode prend en considération l'interaction spatiale dans chaque composante couleur, mais tient compte aussi de l'interaction inter-composantes. L'estimation des paramètres du GMRF est effectuée en utilisant les méthodes d'estimation par maximum de vraisemblance. La segmentation est réalisée à l'aide de deux étapes : la première étape consiste à appliquer récursivement des divisions en régions afin de partitionner l'image en blocs carrés jusqu'à la satisfaction d'un critère d'uniformité. La deuxième étape consiste à effectuer un regroupement par accumulation, ce qui permet de fusionner les régions ayant des caractéristiques similaires. Cette deuxième étape sert à identifier les contours des différentes textures de l'image. PANJWANI et HEALEY affirment que les résultats obtenus sur des images naturelles par utilisation conjointe de la couleur et de la texture sont meilleurs que ceux obtenus avec la couleur ou bien la texture seule.

Une autre méthode de segmentation plus récente, introduite par SHI et FUNT [Shi 2007], se base sur une représentation plus compacte suite à une conversion des composantes couleur de l'espace  $(R, G, B)$  sous la forme de quaternion. En effet, cet outil mathématique (quaternion) est une nouvelle représentation de la couleur qui est dans le cas habituel décrite par trois composantes. Les quaternions permettent de décrire la couleur non plus par un triplet, mais par un seul nombre dit "hypercomplexe" [Shi 2005]. Le quaternion noté  $q$  peut être donné par l'équation 1.5 :

$$q = a + b.i + c.j + d.k \quad (1.5)$$

où  $a$  est la partie réelle du quaternion et  $b, c$  et  $d$  sont trois parties imaginaires. Le triplet de l'espace  $(R, G, B)$  peut être représenté par un quaternion purement imaginaire tel que :

$$q = R.i + G.j + B.k \quad (1.6)$$

Cette technique basée sur la représentation de la couleur par une seule composante au lieu de trois, est réalisée via trois étapes : la première étape consiste à générer des vecteurs d'attributs générés à l'aide d'une "Quaternion principal component analysis" (Q-PCA) sur des données d'apprentissage. La Q-PCA correspond à une analyse en composantes principales appliquée aux quaternions [Bihan 2003]. La deuxième étape consiste à projeter les données d'entrée dans le sous-espace Q-PCA et de classer les vecteurs résultants par l'algorithme de classification des K-moyennes. La dernière étape consiste à appliquer une procédure de fusion de régions adjacentes ayant les mêmes caractéristiques de texture afin d'éviter le problème de sur-segmentation. Les auteurs justifient l'apport de l'utilisation de la technique de quaternion par l'avantage de l'étude simultanée des relations intra et inter-composantes couleur du pixel avec ses voisins. Cependant, cette méthode de segmentation présente l'inconvénient d'être fortement dépendante du choix de la taille des sous-fenêtres permettant l'échantillonnage du contenu couleur/texture local et du choix d'autres paramètres liés à la méthode (le seuil

de fusion de régions et le nombre de classes défini pour le classifieur des K-moyennes).

L'intégration implicite des attributs de couleur et de texture est liée aux premiers travaux d'analyse de la couleur et de la texture. Cette constatation est justifiée en partie par le fait que l'analyse de la texture est principalement évaluée dans le contexte de l'image en niveaux de gris. Par conséquent, l'extension des méthodes de segmentation utilisant la texture monochrome est réalisée en cherchant les attributs de texture dans chaque composante couleur. Par conséquent, l'intégration implicite des deux sources d'information couleur et texture est souvent effectuée par des simples modèles. Contrairement à l'intégration implicite, la deuxième catégorie d'intégration des informations couleur et texture, présentée ci-après, est réalisée grâce à des modèles plus sophistiqués afin d'obtenir des résultats de segmentation plus précis. Dans ce qui suit, nous présentons quelques principaux travaux opérant avec la deuxième catégorie de combinaison couleur/texture.

### 1.3.2/ L'INTÉGRATION SUCCESSIVE

Dans le but de réduire les limites de l'intégration implicite des attributs couleur/texture dans le processus de segmentation, des stratégies alternatives ont été introduites. Parmi ces stratégies d'intégration d'attributs, nous citons l'approche basée sur l'intégration successive qui constitue une direction de recherche populaire. L'idée principale des méthodes faisant partie de cette catégorie consiste à ne pas avoir de règles explicites ou bien de modèles analytiques qui décrivent la dépendance entre la couleur et la texture. Par conséquent, l'extraction de la couleur et de la texture doit se faire en série. À partir de cette conclusion, la segmentation d'images peut être formulée comme étant une approche multi-phases basée sur le partitionnement de l'image du grossier au plus fin (coarse-to-fine). Dans ce qui suit, nous présentons le concept des principales méthodes de segmentation faisant partie de cette catégorie, à savoir les méthodes Jseg, Fusion Clustering Results (FCR), Texture and Boundary Encoding-based Segmentation (TBES) et de segmentation de Felz-Hutt.

#### 1.3.2.1/ JSEG

La méthode de segmentation Jseg, introduite par DENG et MANJUNATH [Deng 2001], consiste à appliquer deux étapes indépendantes pour segmenter une image. La première étape consiste à quantifier la couleur de l'image afin de partitionner l'image en régions représentatives tout en attribuant aux pixels de l'image les valeurs des labels des nouvelles régions de l'image ce qui permet de créer une carte de labels. La deuxième étape, effectuée directement sur la carte de labels, consiste à segmenter l'image en utilisant une méthode de croissance de régions. Dans cette étape, les auteurs définissent un nouveau critère permettant la minimisation d'un coût associé au partitionnement de l'image. Une notion de « J-image » correspondant à des mesures d'homogénéité pour différentes échelles et indiquant l'emplacement des contours potentiels, est également introduite. Par conséquent, la segmentation spatiale est effectuée en appliquant la croissance de régions à partir des germes du « J-image ». Dans le but d'identifier l'homogénéité d'une image, un ensemble d'hypothèses a été défini dans les travaux de DENG et MANJUNATH [Deng 2001] et rappelé ci-dessous.

- chaque image contient un ensemble de régions de texture approximativement homogènes ;
- l'information couleur dans chaque région de l'image peut être représentée par peu de couleurs quantifiées ;
- les couleurs entre deux régions voisines de l'image peuvent être distinguées (c'est l'hypothèse de base de toute méthode de segmentation d'images couleur).

La figure 1.16 illustre le schéma récapitulant les différentes étapes de l'algorithme Jseg. Le lecteur intéressé par des informations supplémentaires sur les différentes étapes illustrées sur la figure 1.16, pourra se référer aux travaux de DENG [Deng 2001].

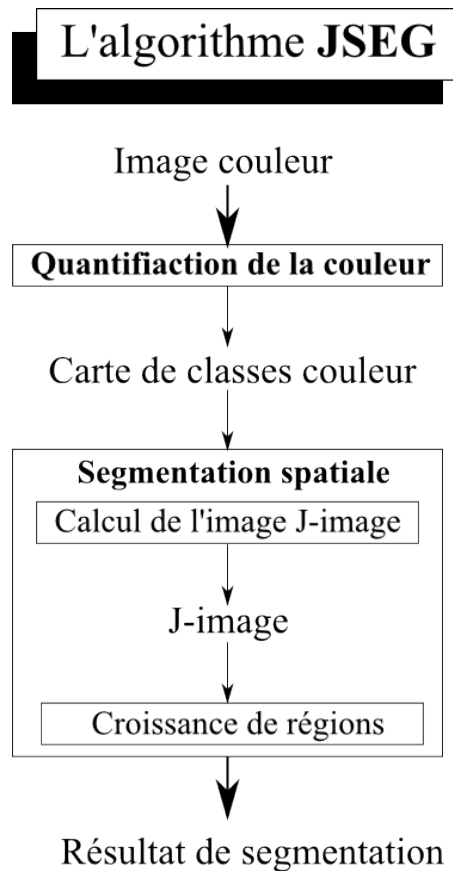


FIGURE 1.16 – Principe de l'algorithme Jseg.

### 1.3.2.2/ FUSION CLUSTERING RESULTS (FCR)

MIGNOTTE propose une nouvelle stratégie de segmentation d'images [Mignotte 2008], en affirmant qu'il est plus simple et efficace de combiner différents modèles de partitions de l'image originale que d'élaborer une nouvelle méthode de segmentation. En effet, la nouvelle stratégie consiste principalement à fusionner des partitions grossières réalisées par l'algorithme des k-moyennes sur l'image originale représentée dans plusieurs espaces couleur. La première étape consiste à réaliser des segmentations grossières. Ces dernières sont, dans une deuxième étape, fusionnées via l'algorithme des k-moyennes en utilisant comme attributs d'entrée les histogrammes locaux des images grossièrement segmentées.

La figure 1.17 et l'algorithme 4 illustrent le principe de cette première étape. Les cartes de segmentation initiales sont obtenues en appliquant l'algorithme de classification des k-moyennes sur l'image originale présentée dans différents espaces couleur et en utilisant des descripteurs de texture multidimensionnels conçus à partir des histogrammes de couleur quantifiée estimés sur le voisinage de chaque pixel à classer. Soient  $b(p) \in \{0, \dots, N_b - 1\}$  l'indice de la case de l'histogramme multidimensionnel,  $(R_p, V_p, B_p)$  le vecteur couleur associé au pixel  $p$  et  $\mathcal{N}_p$  l'ensemble des pixels appartenant à son voisinage.  $h(p) = \{h(n; p)\}_{n=0, \dots, N_b - 1}$  constitue l'ensemble des valeurs des cases du descripteur multidimensionnel caractérisant la distribution de couleur de chaque pixel à classer. Ces valeurs sont calculées à l'aide de l'équation 1.7 donnée ci-après.

$$h(n; p) = \mathcal{K} \sum_{p' \in \mathcal{N}(p)} \delta[b(p') - n] \quad (1.7)$$

où  $\delta$  est la fonction de KRONECKER et  $\mathcal{K} = \frac{1}{(N_w)^2}$  est la constante de normalisation assurant l'égalité  $\sum_{n=0}^{N_b-1} h(n; p) = 1$ .

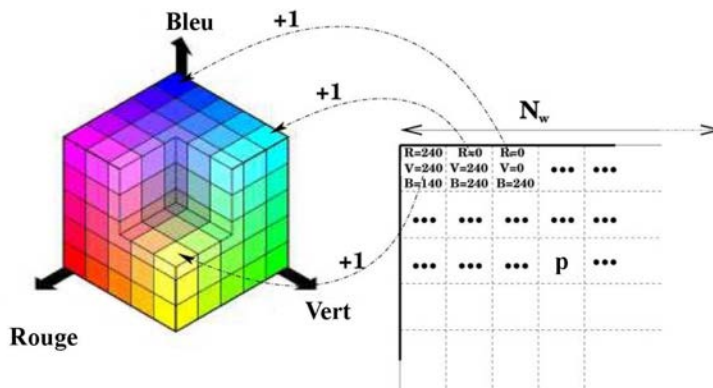


FIGURE 1.17 – Technique de génération d'un descripteur cubique : Pour un pixel  $p$  d'une image dans l'espace couleur  $(R, G, B)$  et dans un voisinage de taille  $N_w \times N_w$  ( $N_w = 7$ ), l'histogramme cubique de taille  $N_b = q^3$  ( $q = 5$ ) est obtenu en incrémentant de (+1) la valeur de la case correspondant aux coordonnées  $R_{p'}$ ,  $V_{p'}$  et  $B_{p'}$  associées au pixel  $p'$  appartenant au voisinage de  $p$  [Mignotte 2008].

---

**Algorithme 4:** Estimation d'un descripteur multidimensionnel local [Mignotte 2008].

---

**Entrées :**  $\mathcal{N}_p$  : l'ensemble des pixels  $\in$  au voisinage de taille  $N_w \times N_w$  et centré au pixel courant  $p$

$\lfloor \rfloor$  désigne la partie entière

**Pour**  $p' \in \mathcal{N}(p)$  *ayant comme information couleur  $R_{p'}$ ,  $V_{p'}$  et  $B_{p'}$*  **faire**

$n \leftarrow q^2 \cdot \lfloor q \cdot R_{p'} / 256 \rfloor + q \cdot \lfloor q \cdot V_{p'} / 256 \rfloor + \lfloor q \cdot B_{p'} / 256 \rfloor$

$h[n] \leftarrow h[n] + \frac{1}{(N_w)^2}$

**fin**

**Sorties :** Estimation des valeurs des  $N_b = q^3$  cases du descripteur

---

Les  $N_b$  cases du descripteur sont ensuite regroupées dans différentes classes via l'algorithme des k-moyennes en utilisant la distance euclidienne. Cette procédure de partitionnement est répétée sur des images obtenues suite à la conversion de l'image

initiale dans différents espaces couleur. Ces images peuvent être vues comme étant des images acquises par différents capteurs. Dans le cadre des travaux [Mignotte 2008],  $N_s = 6$  partitions initiales ont été exploitées, ce qui revient à avoir six espaces couleur différents mis en œuvre à savoir les espaces couleur  $(R, G, B)$ ,  $(H, S, V)$ ,  $(Y, I, Q)$ ,  $(X, Y, Z)$ ,  $(L, A, B)$  et  $(L, U, V)$ .

La deuxième phase de l'approche de MIGNOTTE est la fusion des cartes de partitions. L'idée principale de cette fusion consiste à considérer comme vecteur de descripteurs, pour chaque pixel, l'histogramme local (calculé toujours sur le voisinage centré au pixel) des labels de chaque partition. Par conséquent, pour une fusion de  $N_s$  segmentations (ayant  $K_1$  régions chacune) en une segmentation finale ayant  $K_2$  régions, l'étape préliminaire d'extraction des descripteurs revient à générer  $N_s$  histogrammes (ayant  $K_1$ -cases). Les descripteurs sont ensuite regroupés pour former un vecteur de descripteurs de dimension  $K_1 \times N_s$  (un histogramme multidimensionnel  $(K_1 \times N_s)$ ). La procédure de fusion peut donc être considérée comme un problème de segmentation d'histogrammes locaux des labels des régions des segmentations grossières initiales. La fusion est réalisée encore une fois avec l'algorithme des k-moyennes en exploitant une mesure de similarité entre les histogrammes. La mesure de similarité est obtenue à partir du coefficient de similarité de BHATTACHARYA [Kailath 1967].

### 1.3.2.3/ TEXTURE AND BOUNDARY ENCODING-BASED SEGMENTATION (TBES)

RAO et al. proposent une nouvelle méthode de segmentation d'images [Rao 2009] qui consiste à modéliser les régions homogènes de texture à l'aide d'une distribution gaussienne et à coder les contours de ces régions par une chaîne de code adaptative (FREEMAN chain code). La segmentation optimale obtenue avec cette technique correspond à celle qui utilise les codes les plus courts pour le codage des différentes textures et contours de l'image à analyser. Elle est obtenue via un processus de regroupement hiérarchique ascendant des blocs de l'image de tailles décroissantes. Soient  $\mathcal{I}_{w_M}(R)$  le contenu d'une région  $R$ ,  $\forall R \in \mathcal{R}$ ,  $\Delta L_{w,E}(R_i, R_j)$  représente la différence entre les longueurs codées des régions  $R_i$  et  $R_j$  et leurs contours avant et après la fusion, l'algorithme 5 présenté ci-après, illustre les différentes étapes de la méthode de segmentation TBES.

Notons que pour la méthode de segmentation TBES, une image initialement codée dans l'espace couleur  $(R, G, B)$ , doit être convertie dans l'espace couleur  $(L, a, b)$ . Les valeurs des composantes couleur de cet espace sont ensuite stockées sous forme de vecteurs. Par conséquent, l'intégration de la couleur a été effectuée via cette étape et les attributs de texture sont ensuite extraits en appliquant le codage des régions de texture homogènes et de leurs contours. Pour de plus amples informations sur ce codage, les travaux de RAO [Rao 2009] décrivent la méthode TBES avec plus de détails. Cet algorithme a été testé dans [Rao 2009] sur la base d'images de BERKELEY [Berkeley] et sur la base d'images MSRC de reconnaissance d'objets.

---

**Algorithme 5:** Segmentation basée sur le codage de texture et des contours (TBES) [Rao 2009].

---

**Entrées :**  $I$  image à segmenter,  $w_M$  la taille maximale de la fenêtre,  $\mathcal{R} = \{R_1, \dots, R_k\}$  des superpixels et  $\varepsilon$  une mesure de distorsion.

**Pour**  $w = 2 : w_M$  **faire**

    Stocker dans un vecteur  $X_w$  le contenu de la fenêtre de taille  $w \times w$  centrée sur un pixel  $p$  de  $I$ .

    Construire le vecteur  $\hat{X}_w$  en appliquant l'Analyse des Composantes Principales sur  $X_w$ .

**fin**

Construire RAG (Region Agency Graph)  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  où  $\mathcal{V} \simeq \mathcal{R}$  et  $e_{ij} \in \mathcal{E}$  seulement si  $R_i$  et  $R_j$  sont adjacentes dans  $I$

$w = w_M$

**Répéter**

**Si**  $w = w_M$  **alors**

        Trouver  $R_i$  et  $R_j$  tel que  $e_{ij} \in \mathcal{E}$ ,  $\mathcal{I}_w(R_i) \neq \emptyset$ ,  $\mathcal{I}_w(R_j) \neq \emptyset$  et  $\Delta L_{w,\varepsilon}(R_i, R_j)$  est maximale

**fin**

**Sinon Si** **alors**

        Trouver  $R_i$  et  $R_j$  tel que  $e_{ij} \in \mathcal{E}$ ,  $\mathcal{I}_w(R_i) \neq \emptyset$ ,  $\mathcal{I}_w(R_j) \neq \emptyset$ ,  $\mathcal{I}_{w+2}(R_i) \neq \emptyset$ ,  $\mathcal{I}_{w+2}(R_j) \neq \emptyset$  et  $\Delta L_{w,\varepsilon}(R_i, R_j)$  est maximale

**fin**

**Si**  $\Delta L_{w,\varepsilon}(R_i, R_j) \gtrsim 0$  **alors**

$\mathcal{R} := (\mathcal{R} \setminus \{R_i, R_j\}) \cup \{R_i \cup R_j\}$

        Mettre à jour  $\mathcal{G}$  en utilisant la nouvelle région obtenue après la fusion de régions.

$w = w_M$

**fin**

**Sinon Si**  $w \neq 1$  **alors**

$w = w - 2$

**fin**

**Jusqu'à**  $\mathcal{I}_{w_M}(R) \neq \emptyset$ ,  $\forall R \in \mathcal{R}$  et  $\Delta L_{w_M,\varepsilon}(R_i, R_j) \leq 0$ ,  $\forall R_i, R_j \in \mathcal{R}$ ;

**Sorties :** L'ensemble de régions  $\mathcal{R}$

---

### 1.3.2.4/ SEGMENTATION DE FELZ-HUTT (FH)

FELZENSZWALB et HUTTENLOCHER proposent dans leurs travaux [Felzenszwalb 2004], une méthode de segmentation basée sur la fusion de régions. Cette méthode de segmentation permet de regrouper les pixels d'une image de sorte à ce que la segmentation obtenue soit optimale (ni très grossière ni très fine). Soit un graphe dont les nœuds sont les pixels de l'image et les arêtes reliant ces nœuds sont marqués par des valeurs mesurant la similarité entre ces pixels. Les poids de similarité peuvent prendre, dans le cas le plus simple, la valeur de la différence de couleur. Soit  $Int(R)$  la valeur de la différence interne d'un groupe de nœuds ayant le poids le plus fort. En prenant les arêtes dans l'ordre croissant de leurs poids de similarité, les régions  $R_1$  et  $R_2$  connectées par l'arête courante sont fusionnées si le poids de similarité associé est inférieur au critère défini par l'équation 1.8.

$$\min(Int(R_1) + \tau(R_1), Int(R_2) + \tau(R_2)) \quad (1.8)$$

où  $\tau(R) = \frac{k}{|R|}$  sachant que  $k$  est un paramètre permettant de donner une préférence à la taille d'un groupe de nœuds  $|R|$ .

Il existe d'autres méthodes de segmentation opérant selon l'intégration successive de la couleur et de la texture [Huang 1992, Mirmehdi 2000, Gevers 2002, Scarpa 2006, Wang 2006, Zheng 2006, Yu 2008, Krinidis 2009]. Nous citons quelques unes des principales méthodes sans rentrer trop dans les détails à savoir les travaux de KRINIDIS et PITAS [Krinidis 2009] qui présentent un algorithme nommé MIS (Modal Image Segmentation). Cet algorithme ressemble dans son principe à l'algorithme Jseg [Deng 2001] de DENG et MANJUNATH si l'on remplace l'étape de croissance de régions multi-échelle par une fonction de modèle d'énergie déformable. Cette dernière permet de combiner les intensités des pixels avec l'information spatiale de l'image. Tout comme l'algorithme de Jseg, l'algorithme MIS est composé de deux étapes principales. La première étape implique une représentation grossière de l'image obtenue par application d'une procédure de quantification de la couleur. L'image quantifiée est ensuite utilisée dans une deuxième étape afin d'extraire l'information spatiale en utilisant les forces externes de la fonction d'énergie déformable. Cette étape implique le calcul d'une énergie fonctionnelle mesurant les caractéristiques de la région autour de chaque pixel. L'étape finale de l'algorithme MIS consiste à appliquer une fusion ascendante afin d'éliminer l'effet de sur-segmentation. L'algorithme MIS a été appliqué sur 300 images de la base de BERKELEY [Berkeley] et comparé aux méthodes de segmentation MeanShift [Comaniciu 2002], NCuts [Shi 2000]. Il est important de noter que la performance de l'algorithme MIS est fortement dépendante d'un paramètre de l'équation de la transformée discrète modale [Krinidis 2009].

HEDJAM et MIGNOTTE proposent quant à eux dans leurs travaux [Hedjam 2009], un algorithme de segmentation hiérarchique connu sous le nom de Hierarchical graph-based Markovian Clustering (HMC) qui s'applique en deux étapes. Dans la première étape, l'image initiale est sur-segmentée en  $K$  régions prédéfinies en utilisant les modèles de champs de Markov aléatoires (Markov Random Field Models). Dans la deuxième étape, l'image résultante est modélisée par un Region Adjacency Graph (RAG) classique où une valeur de poids est associée à chaque contour de région permettant de traduire sa similarité aux régions adjacentes. Les régions sont donc identifiées à travers la

liaison des contours et sont représentées par des groupes denses de nœuds fortement connectés.

Bien que cette catégorie d'intégration successive (en série) d'attributs couleur et texture a montré sa performance, son inconvénient majeur réside dans la difficulté de sa généralisation. D'autre part, vu que l'intégration de la couleur et de la texture est réalisée de manière successive, les solutions algorithmiques qui tiennent compte aussi bien de l'homogénéité de la distribution de la couleur et de l'irrégularité de la texture (facteur fort présent dans le cas des images naturelles), sont réduites. De plus, les méthodes de segmentation basées sur l'intégration successive nécessitent l'optimisation d'un large nombre de paramètres, raison qui limite leur application dans les segmentations non-supervisées.

### 1.3.3/ L'INTÉGRATION LORS DU PROCESSUS DE SEGMENTATION

Contrairement aux méthodes de segmentation citées précédemment où les attributs couleur et texture sont calculés en cascade, les méthodes opérant par "intégration lors du processus de segmentation" extraient les attributs de couleur et de texture séparément. Ces méthodes sont basées sur l'hypothèse de modéliser différemment la couleur et la texture lorsqu'elles sont analysées d'un point de vue statistique. L'avantage de cette catégorie d'extraction d'attributs couleur et texture, comparée aux deux autres catégories précédemment introduites, est la possibilité d'optimiser leur contribution lors de leur intégration dans le processus de segmentation. Les méthodes de segmentation faisant partie de ce type d'utilisation conjointe de couleur et de texture peuvent être divisées en deux sous-catégories : la première renferme les méthodes qui combinent les attributs couleur/texture en utilisant des techniques basées sur le traitement des régions à savoir la décomposition/fusion (split/merge) [Chen 2002, Nammalwar 2004, Ilea 2006, Nammalwar 2010], la croissance de région (region growing) [Pappas 1992, Paschos 1999, Rogowitz 2005, Fondon 2006, Griniasa 2008, Ugarriza 2009] et les contours actifs [Freixenet 2004, Allili 2007, Cremers 2007, Garcia 2008, Han 2009, Kim 2009, Brox 2010]. Les méthodes de segmentation basées sur le traitement des régions sont les mécanismes de segmentation les plus étudiés dans le domaine de l'analyse couleur/texture. Ceci est justifié par le fait que la cohérence spatiale entre les pixels adjacents (ou les régions de l'image) est prise en considération lors du processus de segmentation. La deuxième sous-catégorie est définie par les intégrations d'attributs basées sur des techniques statistiques [Tan 1994, Campbell 1996, Jain 1998, Niskanen 2001, Zoller 2002, Khotanzad 2003, Cheng 2003, Martin 2004, Chang 2005, Ilea 2007, Datar 2008, Ilea 2008, Ilea 2009, Ooi 2009, Hanbury 2009] et probabilistes [Manduchi 1999, Jolly 2000, Echigo 2000, Ozyildiz 2002, Kato 2002, Carson 2002, Huawu 2004, Destremes 2006, Xia 2006, Kato 2006, Khan 2009, Fukuda 2009, Serrano 2009]. Dans cette deuxième sous-catégorie, les algorithmes de détection de contours et les méthodes par apprentissage ont également été étudiés dans le cadre de la segmentation par utilisation conjointe de la couleur et de la texture.

Il existe alors plusieurs méthodes de segmentation d'images utilisant l'intégration de couleur et de texture lors du processus de segmentation. Nous citons dans cette section



quelques-unes d'entre-elles. Nous présentons donc deux méthodes de segmentation de la première sous-catégorie à savoir la méthode de segmentation basée sur le modèle de Chan-Vese [Chan 2001, Vese 2002] et la méthode de segmentation basée sur une fonction de variation totale [Donoser 2009]. Nous détaillons également quelques méthodes de segmentation appartenant à la deuxième sous-catégorie basée les stratégies statistiques et probabilistes à savoir la segmentation par combinaison de couleur, texture et luminance [Arbelaez 2011], la segmentation par l'algorithme C-TeX basé sur la cohérence couleur/texture [Ilea 2008], la segmentation par combinaison fixe de gradients couleur et texture [Angulo 2007] et la segmentation par combinaison adaptative ou supremum des gradients couleur/texture [Cohen 2009].

### 1.3.3.1/ SEGMENTATION PAR UNE FONCTION DE VARIATION TOTALE (TOTAL VARIATION SEGMENTATION (TV))

DONOSER et al. proposent dans leurs travaux [Donoser 2009], une nouvelle méthode de segmentation non-supervisée d'images couleur qui consiste à segmenter l'image initiale en plusieurs étapes. En effet, à chaque étape, la segmentation est réalisée sur une partie saillante de l'image. Les résultats intermédiaires obtenus sont ensuite fusionnés pour avoir une seule segmentation composée. L'approche renferme donc trois phases : la première phase concerne la recherche automatique des régions saillantes dans l'image qui peuvent être identifiées en appliquant une segmentation par propagation d'affinité (affinity propagation clustering). Chaque région saillante trouvée servira à définir une initialisation de segmentation indépendante. La deuxième phase consiste à utiliser, de manière indépendante, les résultats de la première phase telle que les régions obtenues sont vues comme des régions de germes pour effectuer une segmentation performante pour chaque région saillante précédemment identifiée. Dans cette phase, un algorithme hybride à base de contours actifs géodésiques (Geodesic Active Contour) [Caselles 1997] est exploité. Cet algorithme utilise un modèle incorporant les informations de couleur et de texture dans une fonction d'énergie convexe basée sur une variation totale pondérée. Ensuite, la segmentation est effectuée par la minimisation de cette fonction d'énergie  $E$  définie par l'équation 1.9 donnée ci-après.

$$\min_C \{E\} = \min_C \left\{ \int_0^{L_C} g(\nabla I(C(s))) ds + \lambda \int_{\Omega} u f d\Omega \right\} \quad (1.9)$$

où  $C$  est le contour de la région étudiée,  $L_C$  est sa longueur euclidienne,  $g$  est la fonction indicatrice du contour qui s'annule en bordure des objets,  $u$  est une fonction caractéristique et  $f$  est la fonction qui englobe l'information sur une région  $\Omega$ .

L'équation est composée de deux termes (deux intégrales), la première est relative au contour de la région qui sera remplacé par une fonction de variation totale pondérée et la deuxième est relative à la région. Pour plus de détails sur cette deuxième étape, le lecteur pourra se référer aux travaux de DONOSER et al. [Donoser 2009]. La troisième phase consiste à fusionner les résultats en une seule segmentation composée en vérifiant l'exactitude des labels locaux permettant de trouver une segmentation globale optimale.

### 1.3.3.2/ SEGMENTATION BASÉE SUR LE MODÈLE CHAN-VESE (CV)

Le modèle Chan-Vese de CHAN et VESE, introduit initialement dans [Chan 2001], a évolué pour être généralisé dans leurs travaux [Vese 2002]. La méthode de segmentation basée sur ce modèle constitue une méthode puissante et flexible capable de segmenter plusieurs types d'images, voire même des images assez difficiles à segmenter par des outils classiques (à savoir les méthodes basées sur le seuillage et les méthodes basées sur les gradients). Généralement, dans le cadre de la détection d'objets, le modèle de contour actif est obtenu en cherchant une partition particulière de l'image en deux régions : une région qui représente l'objet à détecter et l'autre région qui représente l'arrière plan. Le contour actif est donné par la frontière entre les deux régions obtenues. Le modèle obtenu constitue un cas particulier du problème de partition minimale de Mumford-Shah [Mumford 1989] pour la segmentation d'images. Dans les travaux récents de CHAN et VESE, le modèle de contour actif a été généralisé pour être capable de segmenter l'image en plus que deux régions, en proposant un nouveau système multi-phases de "niveau de ligne" (level set) pour le problème de Mumford-Shah [Mumford 1989]. Les auteurs ont montré qu'avec un nombre réduit de fonctions de "niveau de ligne", les topologies complexes peuvent être représentées et les phases de partition ne causent plus les effets de "vide" (vacuum) et de "chevauchement" (overlap). Par exemple, en se basant sur le théorème des quatre couleurs (The Four-Color Theorem), il a été montré que seulement deux fonctions de "niveau de ligne" suffisent théoriquement pour représenter n'importe quelle partition.

### 1.3.3.3/ SEGMENTATION BASÉE SUR LA COHÉRENCE COULEUR/TEXTURE (C-Tex)

L'algorithme C-Tex est une méthode de segmentation récente, introduite par ILEA et WHELAN [Ilea 2008]. Dans cette méthode, la segmentation couleur est un élément majeur du système proposé, impliquant l'analyse statistique des données en utilisant des représentations couleur multi-espaces. La première étape de la segmentation couleur C-Tex consiste à appliquer un filtrage aux données d'entrée en utilisant un algorithme de diffusion isotrope via la technique "Gradient-Boosted Forward and Backward" (GB-FAB) [Ilea 2007]. Cette étape est appliquée d'une part afin d'éliminer le bruit qui peut être présent dans l'image et d'autre part afin d'améliorer la cohérence locale de la couleur. La sélection du nombre de régions ainsi que le choix de leurs centres initiaux constituent les tâches les plus difficiles et doivent être abordées dans l'implémentation des schémas de partitionnement des données statistiques. La méthode de segmentation commence par extraire les couleurs dominantes et identifier le nombre de régions optimal. Ceci est effectué à partir de la première représentation couleur de l'image en utilisant une procédure non-supervisée basée sur des cartes auto adaptatives (Self Organising Map (SOM)). Ensuite, l'algorithme C-Tex consiste à analyser l'image dans un espace couleur complémentaire où le nombre de régions déjà calculé permet d'effectuer la synchronisation entre les deux étapes précédentes. À l'étape finale, les résultats de segmentation obtenus pour chaque espace couleur constituent une entrée pour un processus de segmentation multi-espaces donnant en sortie l'image segmentée finale. La méthode de segmentation C-Tex permet également d'extraire les attributs de texture à partir de l'information de luminance de l'image initiale. Ceci est effectué en utilisant une technique de décomposition à canaux multiples de la texture (multi-channel decomposition of the texture), basée sur les filtres de Gabor. L'intégration des attributs de

couleur et de texture est effectuée à l'aide du classifieur des K-moyennes spatial adaptatif (Adaptive Spatial K-Means (ASKM)) qui divise les cartes de données par échantillonnage adaptatif de la continuité locale de la texture et de la régularité locale de la couleur dans l'image.

#### 1.3.3.4/ SEGMENTATION PAR COMBINAISON DE LUMINANCE, COULEUR ET TEXTURE (CTB)

Parmi les approches récentes de combinaison d'attributs de texture et de couleur, nous pouvons citer les travaux de ARBELAEZ et al [Arbelaez 2011]. MARTIN et al. adoptent dans [Martin 2004], une approche différente pour la détection des contours des objets présents dans l'image et suggèrent une procédure d'apprentissage supervisée à partir de laquelle, ils effectuent une combinaison de la couleur, de la texture et de la luminance de l'image. La première étape de détection des contours consiste à définir une fonction locale notée  $Pb(x, y, \theta)$  permettant de prédire la probabilité d'avoir un contour en un pixel  $p$ , de coordonnées spatiales  $(x, y)$  et dans une orientation donnée  $\theta$ . Cette fonction est construite en calculant un gradient orienté  $G(x, y, \theta)$  à partir de l'intensité de l'image  $I$ . Ce gradient est obtenu en plaçant un bloc de forme circulaire (disque) centré à la position  $(x, y)$ . Ce disque est divisé en deux demi-disques par un diamètre ayant l'orientation  $\theta$ . Pour chaque demi-disque, un histogramme d'intensité locale est généré (cf. figure 1.18). Finalement, la valeur du gradient de magnitude  $G$  à la position  $(x, y)$  est définie par la distance  $\chi^2$  entre les histogrammes  $h_1$  et  $h_2$  des demi-disques (cf. équation 1.10).

$$\chi^2(h_1, h_2) = \frac{1}{2} \sum_i \frac{(h_1(i) - h_2(i))^2}{h_1(i) + h_2(i)} \quad (1.10)$$

Une étape de filtrage est ensuite appliquée avec le filtre de second-ordre Savitzky-Golay [Savitzky 1964] afin de lisser la détection des pics dans la direction orthogonale à  $\theta$ .

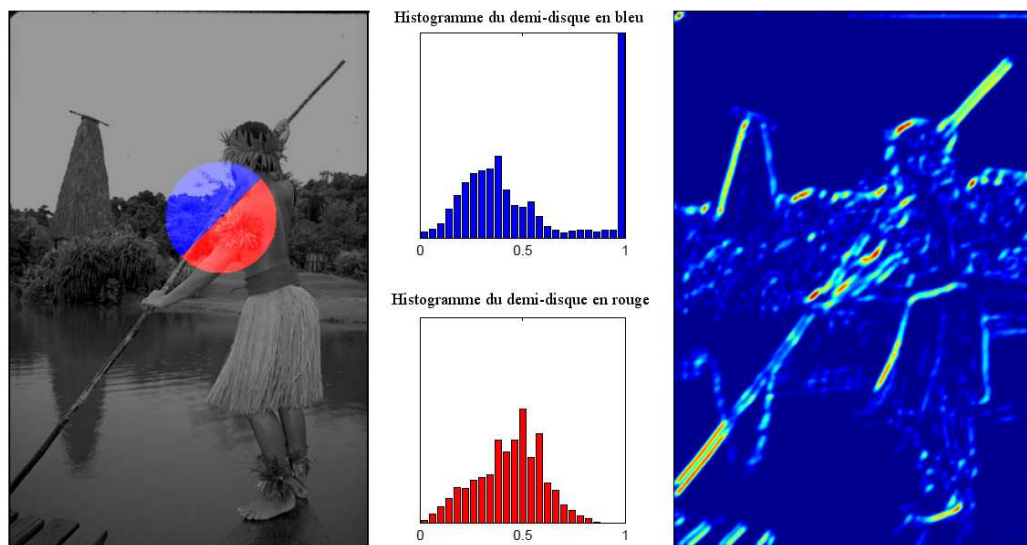


FIGURE 1.18 – Le gradient orienté des histogrammes [Arbelaez 2011]. De gauche à droite : image d'intensité avec un disque circulaire centré en un pixel donné, gradient résultant avec un disque de rayon 5 pixels et une orientation  $\theta = \frac{\pi}{4}$

Le détecteur de contours  $Pb$  permet de combiner les résultats du gradient orienté obtenus à partir de la transformation d'une image d'entrée en quatre canaux séparés en traitant chaque canal séparément. L'image initiale est donc codée dans l'espace couleur  $(L, a, b)$  et les différents canaux à considérer sont : la luminance donnée par la première composante couleur du  $(L, a, b)$ , la composante couleur  $a$  et la composante couleur  $b$  de ce même espace. La quatrième source d'information est donnée par le canal de texture obtenue par la procédure suivante : un texton  $id$  calculé à partir de l'étape de filtrage, est attribué à chaque pixel. L'étape de filtrage consiste à convertir l'image initiale en une image en niveaux de gris et à appliquer ensuite un banc de 17 filtres Gaussien (cf.figure 1.19) [Arbelaez 2011]. Par conséquent, un vecteur contenant les 17 réponses est associé à chaque pixel. Les vecteurs résultants sont par la suite classés par l'algorithme des K-moyennes. Les centres des classes définissent un ensemble d'images spécifiques de textons. Chaque pixel est ensuite associé à un entier  $id \in [1, K]$  du centre de classe le plus proche.

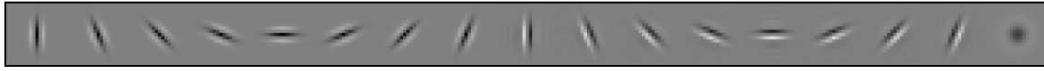


FIGURE 1.19 – Les 17 filtres gaussiens [Arbelaez 2011].

Tout comme pour les canaux de luminosité et de couleurs, les différences d'histogrammes sont calculées sur l'image résultante ayant des valeurs de pixels appartenant à l'intervalle  $[1, K]$ . Par conséquent, en combinant tous les descripteurs de l'image déjà calculés (gradients couleur, gradient texture et gradient de luminance), le détecteur de contours multi-échelle pour une position  $(x, y)$  donnée et dans une orientation  $\theta$  donnée, est obtenu par l'équation 1.11 :

$$mPb(x, y, \theta) = \sum_s \sum_i \alpha_{i,s} G_{i,\sigma}(x, y, \theta) \quad (1.11)$$

où  $s$  indique l'échelle,  $i$  désigne un descripteur parmi la luminance, la "couleur  $a$ ", la "couleur  $b$ " ou bien la texture et  $G_{i,\sigma}(x, y, \theta)$  indique la mesure de la différence d'histogrammes dans le canal  $i$  entre deux demi-disques de rayon  $\sigma(i, s)$  centrés à la position  $(x, y)$  et divisés suivant l'angle  $\theta$ . Le paramètre  $\alpha_{i,s}$  représente le poids relatif à la contribution de chaque gradient de magnitude.

Les auteurs ARBELAEZ et al. définissent 8 valeurs de  $\theta \in [0, \pi]$  et prennent la direction donnant la valeur maximale de la fonction  $mPb$  (cf. équation 1.12).

$$mPb(x, y) = \max_{\theta} \{mPb(x, y, \theta)\} \quad (1.12)$$

Une étape supplémentaire de globalisation basée sur la classification spectrale est effectuée. Elle permet de donner une nouvelle expression de la probabilité de détection de contours globalisée  $gPb$ , définie en fonction d'une somme pondérée de l'information locale et de l'information spectrale, comme le montre l'équation 1.13.

$$gPb(x, y, \theta) = \sum_s \sum_i \beta_{i,s} G_{i,\sigma}(x, y, \theta) + \gamma \cdot sPb(x, y, \theta) \quad (1.13)$$

Tout comme dans l'expression du détecteur de contours multi-échelle  $mPb$ , les poids  $\beta_{i,s}$  et  $\gamma$  sont obtenus par un apprentissage effectué à l'aide de la base d'apprentissage (200 images) de BERKELEY [Martin 2001].

Les résultats de détection de contours des objets présents dans l'image sont ensuite injectés dans la procédure de segmentation introduite par ARBELAEZ et al. [Arbelaez 2011]. Étant donné que nous nous intéressons dans cette section qu'aux méthodes d'intégration de couleur et de texture, nous nous contentons de rappeler brièvement les principales étapes abordées dans cette méthode de segmentation et nous laisserons le lecteur souhaitant plus de détails se référer à [Arbelaez 2011].

Les auteurs introduisent une nouvelle transformation de Ligne de Partage des Eaux orientées (Oriented Watershed Transform (OWT)) afin de produire, de manière hiérarchique, un ensemble initial de régions à partir des résultats obtenus à l'étape de détection de contours. Une carte ultramétrique de contours (Ultrametric Contour Map (UCM)) [Mumford 1989], est ensuite construite, à partir des bordures de ces régions initiales. Cette séquence d'opérations (OWT-UCM) peut être vue comme une machine générique permettant de passer des contours à une carte de régions. Les contours encodés dans la segmentation hiérarchique résultante conservent les valeurs réelles des poids indiquant la probabilité d'être de véritables frontières. Le résultat final est un ensemble de contours fermés pouvant être traités comme un résultat de segmentation ou bien comme un résultat de détection de contours.

### 1.3.3.5/ SEGMENTATION PAR COMBINAISON COULEUR/TEXTURE FIXE (CF)

ANGULO propose dans ces travaux [Angulo 2007], une méthode de segmentation combinant les informations de couleur et de texture. Les attributs de texture sont déterminés par des opérateurs de la morphologie mathématique afin de générer des gradients décrivant la texture. L'information de couleur est issue quand à elle d'un gradient morphologie couleur. La méthode de segmentation par combinaison couleur/texture fixe utilise conjointement la couleur et la texture pour générer un gradient structurel, ce qui permet de réaliser une segmentation structurelle. Cette approche consiste à décomposer tout d'abord l'image en deux composantes définies par une couche de texture et une couche des objets. La couche des objets obtenue en simplifiant l'image originale, sert à calculer le gradient couleur  $G_{col}$ . La couche de texture est générée en cherchant le résidu des composantes de luminance des images originale et simplifiée. Une analyse multi-échelle locale est ensuite appliquée sur la couche de texture en utilisant des opérateurs morphologiques à savoir l'ouverture, la fermeture et des nivellements sur des Filtres Alternés Séquentiels (ASF). Le gradient de texture est noté  $G_{tex}$ . La définition du gradient structurel qui met en jeu un paramètre  $\alpha$  est donnée par l'équation 1.14. Bien que cette approche donne de bons résultats, elle reste une méthode paramétrique qui nécessite la définition de la valeur optimale du paramètre de pondération  $\alpha$  pour chaque image.

$$G_{struc}(I) = (1 - \alpha)G_{tex}(I) + \alpha G_{col}(I) \quad (1.14)$$

### 1.3.3.6/ SEGMENTATION PAR COMBINAISON COULEUR/TEXTURE ADAPTATIVE OU SUPREMUM (CA, CS)

Dans ce même contexte de segmentation par combinaison couleur/texture hybride, COHEN et al. proposent une méthode de combinaison adaptative permettant de s'affroncher du problème du choix du paramètre de pondération  $\alpha$  intervenant dans la combinaison d'attributs couleur et texture. En effet, le paramètre global  $\alpha$ , précédemment défini dans les travaux d'ANGULO, ne prend pas en compte la variation locale des informations de couleur et de texture (le contenu de l'image). Cet inconvénient nécessite alors d'ajuster la valeur du paramètre de pondération  $\alpha$  en fonction des images à traiter afin d'obtenir des résultats satisfaisants. La méthode de combinaison couleur/texture proposée par COHEN et al. consiste à combiner deux gradients (couleur et texture), en s'adaptant localement au contenu de l'image. Cette méthode a été testée sur deux applications différentes [Cohen 2009, Meurie 2010b], avec un gradient morphologie<sup>6</sup> couleur (calculé à l'aide de l'équation 1.15) et le même gradient texture que celui défini dans l'approche d'ANGULO (calculé à partir d'un Filtre Alterné Séquentiel (ASF)) [Angulo 2007]. D'autres gradients de texture basés sur les filtres de Gabor et les descripteurs d'HARALICK ont également été testés pour une meilleure adaptation à la variation locale de l'information de texture [Cohen 2010a].

$$G_{col}(I) = \delta_B(I) - \epsilon_B(I) \quad (1.15)$$

où  $B$  est l'élément structurant utilisé.

Le gradient morphologique texture a été calculé à partir des descripteurs texture calculé de manière locale. Soit  $d_{\vec{r}}$  un descripteur de texture quelconque. Le gradient de texture  $G_{tex}$  d'une image initiale couleur  $I$  est le supremum des gradients morphologiques des images descripteurs comme exprimé par l'équation 1.16 donnée ci-après.

$$G_{tex}(I) = \bigvee_{\vec{r} \in \mathcal{T}} [\delta_B(d_{\vec{r}}) - \epsilon_B(d_{\vec{r}})] \quad (1.16)$$

où  $\mathcal{T}$  est l'ensemble de vecteurs indiquant le voisinage considéré pour le calcul des matrices de co-occurrences.  $\delta_B$  et  $\epsilon_B$  sont respectivement la dilatation morphologie et l'érosion morphologie avec un élément structurant  $B$ .

Le gradient couleur/texture final est un gradient structurel. Ce dernier est ensuite utilisé dans l'algorithme de segmentation par Ligne de Partage des Eaux (LPE). Le gradient couleur est défini sur trois composantes couleur, tandis que le gradient texture, généré à partir d'une image en niveaux de gris, est défini sur une seule composante. Par conséquent, pour combiner ces deux gradients, les auteurs décomposent le gradient couleur en trois composantes  $G_{col}^R$ ,  $G_{col}^G$ ,  $G_{col}^B$  et combinent chacun des gradients obtenus avec le gradient texture selon l'équation 1.17 (cf. figure 1.20).

---

6. Le résidu de la dilatation et de l'érosion de l'image.

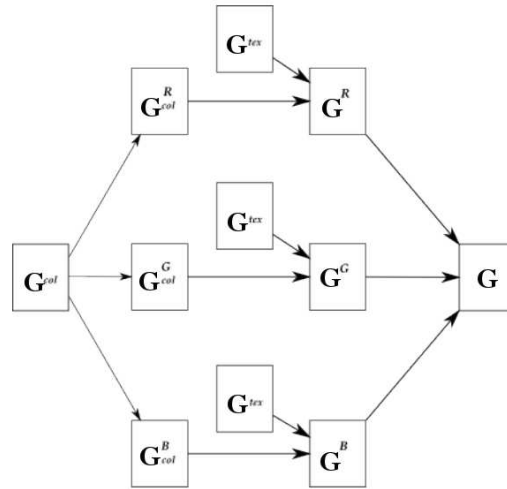


FIGURE 1.20 – Formation du gradient composite par fusion des gradients couleur et texture [Cohen 2010a].

$$\begin{cases} G^R = G_{col}^R \otimes G_{tex} \\ G^G = G_{col}^G \otimes G_{tex} \\ G^B = G_{col}^B \otimes G_{tex} \end{cases} \quad (1.17)$$

Contrairement à la combinaison fixe, la combinaison adaptative utilise un coefficient de pondération  $\alpha_p \in [0, 1]$  adaptatif et calculé automatiquement pour chaque pixel  $p$  de l'image  $I$ , afin de donner (pour chaque pixel) un poids plus important à l'information couleur si celle-ci est dominante par rapport à la texture et inversement. Soient  $t$  et  $r$  deux signaux désignant respectivement la texture et la couleur, le paramètre de pondération adaptatif est donc donné par l'équation 1.18 :

$$\alpha_p = \frac{t(p)}{t(p) + r(p)} \quad (1.18)$$

Les auteurs proposent également une approche de combinaison couleur/texture supplémentaire, appelée combinaison supremum, qui consiste à prendre le maximum de l'information locale (couleur ou texture). Finalement, la segmentation est effectuée via l'algorithme de LPE (utilisant en entrée deux images : une image de gradient et donc le gradient structurel calculé et une image de germes).

## 1.4/ CONCLUSION

Nous avons dans ce chapitre, passé en revue un certain nombre de méthodes de segmentation d'images assez connues de la littérature, utilisant conjointement les informations de couleur et de texture. Celles-ci ont été classées en trois catégories principales à savoir une combinaison couleur/texture par intégration implicite, par intégration successive et appliquée lors du processus de segmentation. Chacune de ces méthodes comporte bien évidemment ses propres avantages et inconvénients que nous avons tentés de synthétiser dans le tableau 1.1 donné ci-après, et ce afin de positionner notre méthode face à la littérature.

Catégorie d'intégration	Méthode de segmentation	Critères		
		Seuillage et paramétrage	Complexité	Nouveauté
Implicite	CTM	+	++	++
	MeanShift	+	+	+
	NCuts	-	++	+
Successive	Jseg	-	+	+
	FCR	+	++	++
	TBES	+	++	+++
	FH	-	+	+
Effectuée lors de la segmentation	TV	-	++	+++
	CV	-	+	+
	C-Tex	-	++	++
	CTB	+	++	+++
	CF	+	+	++
	CA, CS	+	+	+++

TABLE 1.1 – Avantages et limites des principales méthodes de segmentation d'images de la littérature utilisant conjointement les informations de couleur et de texture.

La communauté en vision par ordinateur s'intéresse fortement à la dernière catégorie d'intégration (effectuée lors de l'étape de segmentation) qui se voit prometteuse au vue de la qualité des segmentations obtenues par les méthodes récentes de cette catégorie. C'est la raison pour laquelle, nous nous sommes intéressé à cette catégorie d'intégration et avons proposé dans la suite de ce manuscrit une méthode de combinaison couleur/texte adaptative et non paramétrique réalisant l'intégration couleur/texte lors du processus de segmentation par croissance de régions. Avant de détailler la méthode de combinaison proposée dans cette thèse, il nous a semblé essentiel de nous intéresser à l'évaluation des algorithmes de segmentation combinant ces deux informations. En effet, comme l'illustrent les figures 1.21 et 1.22, les méthodes de segmentation effectuant l'intégration des informations de couleur et de texture lors de la phase de segmentation semblent visuellement offrir de meilleurs résultats. Cependant, cela restant subjectif, nous avons choisi de présenter dans le chapitre suivant les différentes méthodes et métriques d'évaluation permettant de quantifier précisément les performances de chaque méthode de segmentation, sur deux bases d'images.





FIGURE 1.21 – Résultats des principales méthodes de segmentation sur une sélection d'images de la base de BERKLEY [Berkeley] (de haut en bas : images originales, méthode de segmentation CTM [Yang 2008], méthode de segmentation MeanShift [Comaniciu 2002], méthode de segmentation Jseg [Deng 2001] et méthode de segmentation proposée par FELZENSZWALB et HUTTENLOCHER [Felzenszwalb 2004]).



FIGURE 1.22 – Résultats des principales méthodes de segmentation sur une sélection d'images de la base de BERKLEY [Berkeley] (de haut en bas : images originales, méthode de segmentation proposée par CHAN et VESE [Chan 2001], méthode de segmentation par combinaison couleur, texture et luminance [Arbelaez 2011], méthode de segmentation par combinaison couleur/texture fixe [Angulo 2007] et méthode de segmentation par combinaison couleur/texture adaptative [Cohen 2009]).

# ÉVALUATION DE LA SEGMENTATION ET BASES DE TEST

## Sommaire

<b>2.1</b>	<b>Introduction</b>	<b>57</b>
<b>2.2</b>	<b>Les méthodes d'évaluation de la segmentation</b>	<b>58</b>
2.2.1	Les différentes classifications des méthodes d'évaluation	58
2.2.2	Méthodes d'évaluation supervisée	61
2.2.3	Méthodes d'évaluation non-supervisée	70
<b>2.3</b>	<b>Les bases de tests/apprentissage/validation</b>	<b>74</b>
2.3.1	Bases d'images couleur	75
2.3.2	Bases d'images texture	77
<b>2.4</b>	<b>Comparatif des méthodes de segmentation couleur/texture</b>	<b>79</b>
2.4.1	Choix de bases de tests	79
2.4.2	Choix et développement des métriques d'évaluation	80
2.4.3	Expérimentation et discussions	82
<b>2.5</b>	<b>Conclusion</b>	<b>96</b>

## 2.1/ INTRODUCTION

Toute étude dédiée à l'évaluation de la segmentation d'images se manifeste par le développement d'un ensemble de techniques permettant de mesurer la performance de l'algorithme de segmentation. Pour certaines applications, l'algorithme de segmentation est considéré performant s'il reprend parfaitement le système perceptuel humain alors que pour d'autres applications, l'efficacité du calcul ou bien la stabilité de la méthode sont prioritaires par rapport à la performance visuelle. Comme mentionné dans le chapitre précédent, il existe de nombreuses méthodes de segmentation d'images, ce qui rend la tâche d'évaluation fondamentale pour pouvoir comparer ces algorithmes. En effet, l'évaluation est capitale puisqu'elle permet aux utilisateurs des méthodes de segmentation d'images d'améliorer les méthodes de segmentation existantes et de proposer de nouvelles approches plus robustes et/ou adaptées à certaines applications. Bien que l'évaluation de la segmentation attire de plus en plus l'attention des chercheurs [Jiang 2006, Ge 2007, Zhang 2008], elle reste un domaine de recherche émergent qui suscite moins d'intérêt que la segmentation elle-même ; mais ce domaine doit être

considéré afin de toujours disposer d'une méthode adaptée à l'objectif visé (cf. figure 2.1).

Dans ce chapitre, nous avons fait le choix de classer les méthodes d'évaluation en deux catégories : sans et avec segmentation de référence (vérité-terrain). Dans un premier temps, nous présentons quelques méthodes d'évaluation de la littérature. Ensuite, nous évaluons à l'aide de ces méthodes et sur une sélection de bases d'images de la littérature, les différents algorithmes de segmentation d'images décrits dans le chapitre précédent.

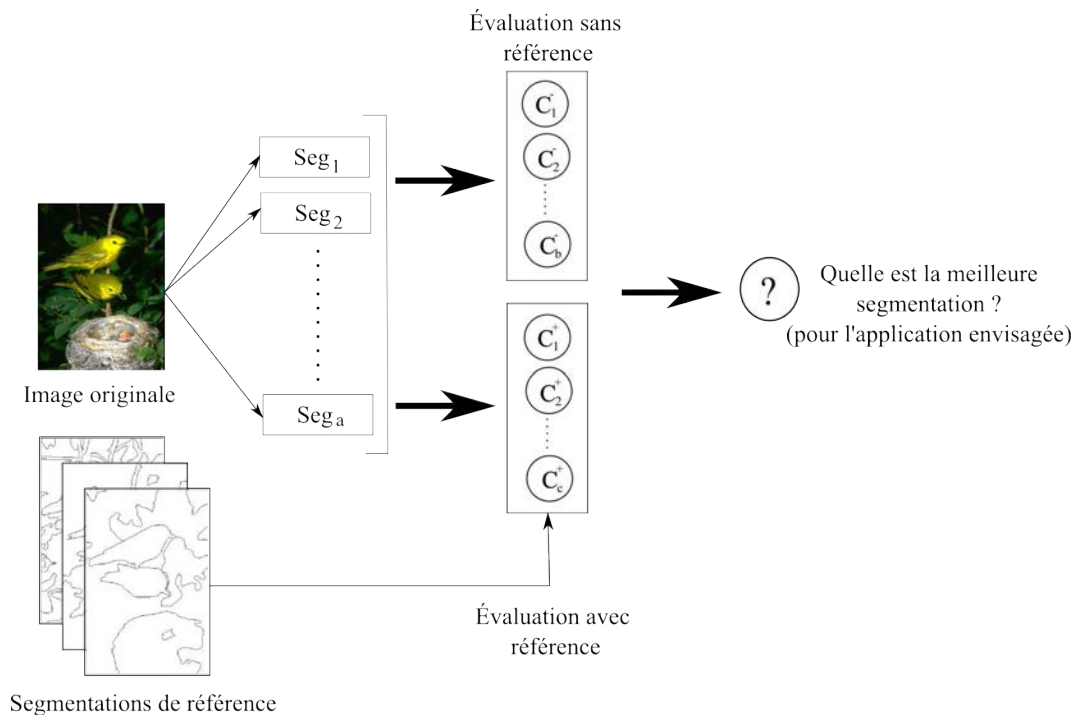


FIGURE 2.1 – Quelle segmentation et quelle métrique d'évaluation choisir ?

## 2.2/ LES MÉTHODES D'ÉVALUATION DE LA SEGMENTATION

Dans cette section, nous dévoilons tout d'abord les différentes taxonomies des techniques d'évaluation comme elles ont été introduites dans la littérature. Ensuite, nous présentons les deux catégories de métriques d'évaluation de la segmentation d'images : les méthodes d'évaluation supervisée (avec vérité-terrain) et les méthodes d'évaluation non-supervisée (sans vérité-terrain).

### 2.2.1/ LES DIFFÉRENTES CLASSIFICATIONS DES MÉTHODES D'ÉVALUATION

Dans la littérature, les méthodes d'évaluation ont été classées selon plusieurs taxonomies en fonction de ce que doit être évalué [Zhang 2008, Jiang 2006, Zhang 1996]. Ces classifications convergent généralement vers les mêmes concepts mais diffèrent en terminologie. Par exemple, ZHANG [Zhang 1996] regroupe les différentes techniques d'évaluation en deux classes à savoir, les méthodes analytiques et les méthodes empiriques (cf. figure 2.2) :



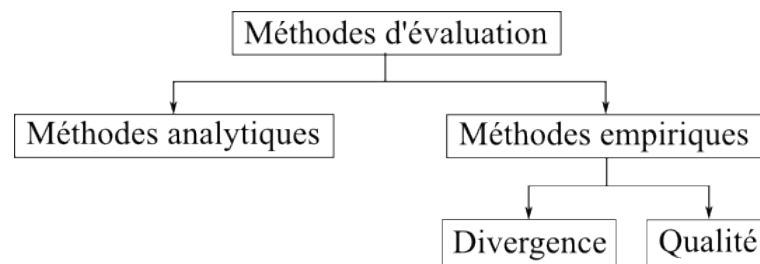


FIGURE 2.2 – Classification des méthodes d'évaluation proposée par ZHANG [Zhang 1996].

- Méthodes analytiques : ces méthodes évaluent la qualité d'un algorithme de segmentation en analysant ses propriétés mathématiques. Elles ne sont généralement pas génériques puisqu'elles ont été définies et adaptées pour des applications précises.
- Méthodes empiriques : ces méthodes peuvent être classées en deux sous catégories : la première concerne les méthodes empiriques de divergence (empirical discrepancy methods) qui consistent à calculer une mesure d'accord entre le résultat de segmentation et une segmentation de référence. La deuxième sous catégorie concerne les méthodes empiriques de qualité (empirical goodness methods) qui consistent à calculer la performance d'une méthode de segmentation en examinant si les résultats vérifient certaines considérations souhaitées.

JIANG et al. [Jiang 2006] quant à eux proposent une répartition différente des méthodes d'évaluation de segmentation d'images. Selon eux, les méthodes d'évaluation peuvent être classées en deux catégories : l'une théorique, l'autre expérimentale (cf. figure 2.3) :

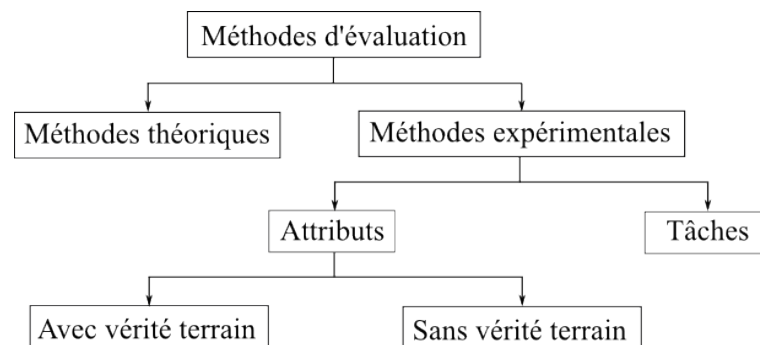


FIGURE 2.3 – Classification des méthodes d'évaluation proposée par JIANG et al. [Jiang 2006].

- Évaluation théorique : elle correspond généralement aux méthodes analytiques que nous avons décrites ci-dessus.
- Évaluation expérimentale : elle inclut les méthodes dites "task-based" qui évaluent la segmentation dans le contexte d'une application. Elle inclut également les méthodes qui englobent deux sous-catégories de méthodes : les méthodes avec vérité-terrain et les méthodes sans vérité-terrain, ce qui revient grossièrement à la classification des méthodes empiriques proposée par ZHANG.

D'autres auteurs [Yang 1995, Chabrier 2006, Zhang 2008] proposent des taxonomies similaires en utilisant de nouvelles terminologies comme le terme supervisé (méthodes

qui utilisent une image vérité-terrain) et le terme non-supervisé (méthodes qui n'utilisent pas d'image vérité-terrain). De manière générale, les schémas de classification des méthodes d'évaluation sont souvent redondants et chaque proposition a ses limites. Pour cette raison, MCGUINNESS [McGuinness 2009] propose un nouveau schéma plus complet dans lequel les méthodes d'évaluation sont classées selon trois rubriques : *objectif*, *référence* et *cible* (cf. figure 2.4).

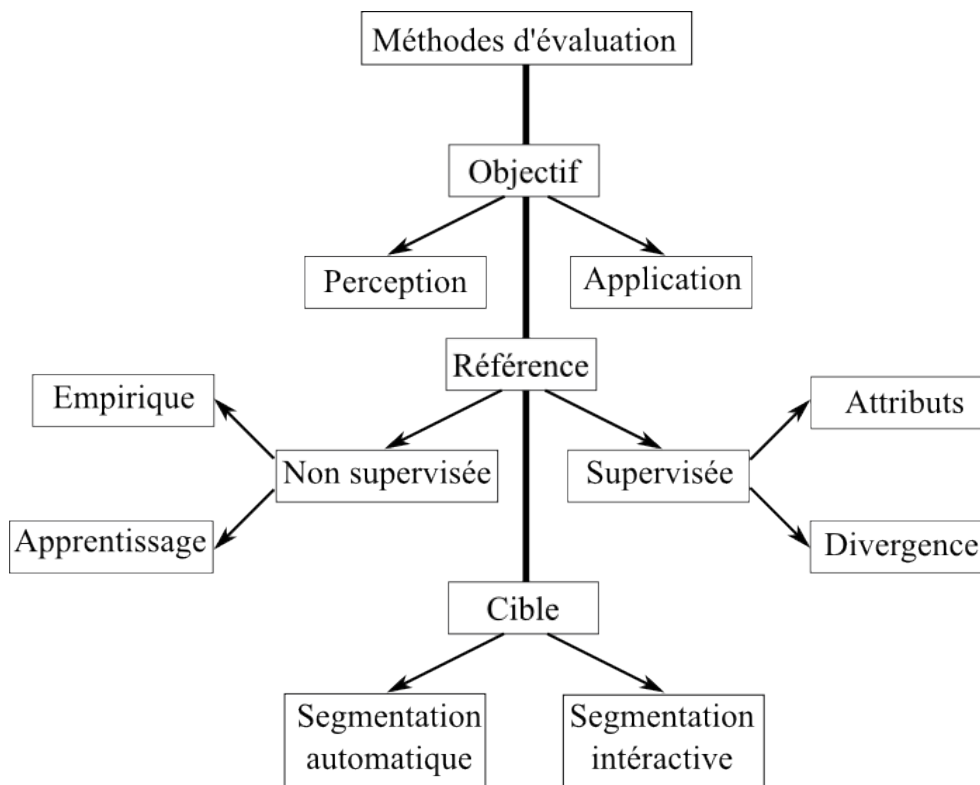


FIGURE 2.4 – Classification des méthodes d'évaluation proposée par MCGUINNESS [McGuinness 2009].

La rubrique *objectif* est utilisée afin de spécifier ce que la méthode d'évaluation vise à évaluer. Dans cette rubrique, deux familles de méthodes d'évaluation ont été définies : la première (perception-based) est dédiée à l'évaluation de l'approximation de la méthode de segmentation par rapport à la segmentation de la perception humaine. La deuxième famille (application-based) s'adresse à l'application pour laquelle la méthode de segmentation est utilisée. Par conséquent, les méthodes d'évaluation appartenant à cette famille mesurent la performance et la pertinence de l'algorithme de segmentation pour un domaine d'application particulier.

La seconde rubrique définie par MCGUINNESS est la *référence*. Elle caractérise les méthodes d'évaluation utilisant une vérité-terrain. Dans cette rubrique, toute évaluation utilisant une ou plusieurs segmentations de référence est considérée comme supervisée et toute évaluation qui ne requiert pas de segmentation de référence est considérée comme non supervisée. Le schéma introduit par JIANG et al. [Jiang 2006] met en évidence l'hypothèse implicite suivante : 1/ les algorithmes d'évaluation application-based n'utilisent pas de vérité-terrain. MCGUINNESS considère que cette hypothèse

est fausse et suppose que les méthodes d'évaluation non supervisée nécessitent une étape d'apprentissage. Il propose pour ces méthodes une subdivision en deux sous-groupes : les méthodes avec apprentissage et les méthodes empiriques. Les méthodes d'évaluation avec apprentissage visent à apprendre automatiquement les propriétés permettant de juger si un résultat de segmentation est acceptable à partir d'un ensemble de données d'apprentissage. À l'inverse, les méthodes d'évaluation purement empiriques spécifient explicitement la performance d'une segmentation à partir des résultats obtenus sans recours à une phase d'apprentissage.

Tout comme les méthodes d'évaluation non supervisée, les méthodes d'évaluation supervisée sont divisées en deux sous-groupes : une évaluation basée sur les attributs (feature-based) et une évaluation basée sur la divergence (discrepancy-based). Cette dernière consiste à calculer une mesure de disparité entre l'image segmentée et l'image de référence. L'évaluation basée sur les attributs consiste à extraire un ensemble d'attributs de l'image segmentée et à les comparer avec les mêmes attributs extraits de l'image de la segmentation de référence. Notons que l'évaluation empirique de qualité (respectivement de divergence) introduite par ZHANG [Zhang 1996] correspond à l'évaluation non supervisée empirique (respectivement à l'évaluation supervisée de divergence) dans le schéma proposé par MCGUINNESS.

La dernière rubrique dans la classification des méthodes d'évaluation proposée par MCGUINNESS est la *cible*. Elle indique le type de l'algorithme de segmentation à évaluer, vu qu'une évaluation d'une segmentation automatique ne peut être pas directement appliquée à une segmentation interactive sachant que cette dernière nécessite forcément l'intervention humaine.

Maintenant que nous avons décrit les grands schémas de classification des méthodes d'évaluation de segmentation d'images cités dans la littérature, nous allons présenter dans les deux sections suivantes les méthodes d'évaluation supervisée et non supervisée les plus populaires et fréquemment utilisées.

### 2.2.2/ MÉTHODES D'ÉVALUATION SUPERVISÉE

Les méthodes d'évaluation supervisée mesurent la performance d'un algorithme de segmentation d'images en comparant l'image segmentée automatiquement avec une image de référence ou vérité-terrain (c-à-d segmentée manuellement par un/plusieurs experts). En effet, le degré de similarité entre la segmentation de référence et la segmentation automatique (obtenue via l'algorithme de segmentation) permet d'évaluer la qualité de la méthode de segmentation. L'intérêt potentiel d'une évaluation supervisée réside dans le fait que la comparaison entre la segmentation automatique et la segmentation de référence est censée fournir une évaluation très précise. En revanche, son inconvénient majeur réside dans la production de la segmentation de référence. En effet, la création (manuelle) d'une vérité-terrain reste une tâche difficile, subjective et lente. De plus, il est généralement difficile de juger qu'une segmentation de référence réalisée par un expert est meilleure qu'une autre. Ceci rend délicat le choix de la vérité-terrain à considérer pour comparer/valider une approche de segmentation d'images lorsque plusieurs experts ont fait l'exercice de créer une segmentation de référence. À

titre d'exemple, nous montrons dans la figure 2.5 plusieurs segmentations de référence pour une image de la base de BERKELEY [Berkeley] afin de mettre en avant cette difficulté.

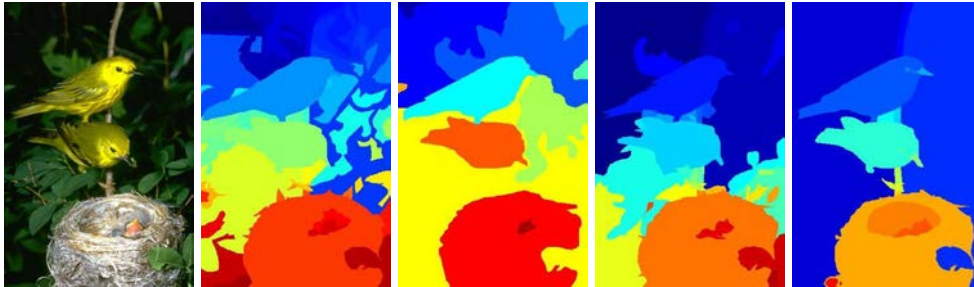


FIGURE 2.5 – Illustration de plusieurs segmentations de référence réalisées pour une même image.

Nous allons maintenant présenter trois premières métriques d'évaluation fréquemment utilisées dans le cadre d'une évaluation supervisée. Elles consistent à traiter la segmentation d'images comme un problème de classification [Jiang 2006]. Pour cela, nous introduisons quelques notations à suivre. Soit  $O = \{p_1, \dots, p_n\}$  l'ensemble d'objets à classer (un objet correspond à un pixel) en un ensemble de groupes disjoints noté  $S = \{r_1, \dots, r_k\}$ . Pour évaluer la performance d'une segmentation automatique par rapport à une segmentation de référence, les deux images segmentées sont d'abord modélisées par les deux ensembles  $S = \{r_1, \dots, r_k\}$  et  $VT = \{r_1, \dots, r_l\}$  puis comparées. Cette notation étant dédiée initialement à la classification, nous modifions la terminologie de la segmentation par celle de la classification afin de définir les trois premières métriques d'évaluation. Par conséquent, nous désignons les ensembles  $S$  et  $VT$  comme étant deux classifications à comparer. Plusieurs mesures de distance ont été définies, et celle qui a été retenue consiste à compter le nombre de paires d'objets  $(p_i, p_j)$  appartenant soit à la même classe soit à des classes différentes. Quatre ensembles peuvent être donc créés en fonction de la disposition et de l'appartenance des paires d'objets  $(p_i, p_j)$  à une classe :

- $T_{11} = \{(p_i, p_j) \text{ tel que } p_i, p_j \text{ appartiennent à la même classe dans } S \text{ et } VT\}$
- $T_{10} = \{(p_i, p_j) \text{ tel que } p_i, p_j \text{ appartiennent à la même classe dans } S \text{ mais pas dans } VT\}$
- $T_{01} = \{(p_i, p_j) \text{ tel que } p_i, p_j \text{ appartiennent à la même classe dans } VT \text{ mais pas dans } S\}$
- $T_{00} = \{(p_i, p_j) \text{ tel que } p_i, p_j \text{ appartiennent à des classes différentes dans } S \text{ et } VT\}$

Soit  $n_{xx} = |T_{xx}|$  le cardinal de l'ensemble  $T_{xx}$  avec  $x \in \{0, 1\}$  et  $n$  le nombre total d'objets (pixels) dans l'image. La relation entre les quatre ensembles est donnée par l'équation suivante :

$$n_{00} + n_{10} + n_{01} + n_{11} = \frac{n(n-1)}{2} \quad (2.1)$$

Dans le cas applicatif, les cardinaux des quatre ensembles sont calculés à partir d'une matrice de confusion [Stehman 1997]. Soient deux classifications/segmentations  $S$  et  $VT$  ayant respectivement  $l$  et  $k$  classes/régions, la matrice de confusion notée  $M$  est donc une matrice de taille  $k \times l$  (cf. figure 2.6) tel que l'élément  $m_{ij}$  est la surface de l'intersection entre les classes/régions d'indice  $i$  dans  $S$  et d'indice  $j$  dans  $VT$ .



		Vérité terrain (VT)		
Segmentation (S)	$r_l$		$r_j$	$r_i$
	$r_l$			
	$r_i$		$m_{ij}$	
	$r_k$			
	$r_k$			

FIGURE 2.6 – La matrice de confusion définie pour deux classifications  $S$  et  $VT$ .

En utilisant les définitions des quatre ensembles décrits ci-dessus, trois mesures d'erreur peuvent être déduites : la mesure de RAND, la mesure de JACCARD et la mesure de FOWLKES et MALLOWS.

### 2.2.2.1/ LA MESURE DE RAND

La mesure de RAND, notée  $\mathcal{R}$ , a été introduite en 1971 par RAND [Rand 1971]. Cette métrique, qui a été définie pour l'évaluation d'une classification, consiste à calculer la somme des paires de pixels ayant le même label dans les deux classifications/segmentations (automatique et de référence) divisée par le nombre total de paires de pixels dans l'image (cf. équation 2.2).

$$\mathcal{R}(S, VT) = 1 - \frac{n_{11} + n_{00}}{\frac{n(n-1)}{2}} \quad (2.2)$$

Les valeurs de  $\mathcal{R}$  sont comprises entre 0 et 1. Une valeur proche de 0 indique une classification/segmentation précise et à l'inverse une valeur proche de 1 indique une erreur de classification/segmentation plus importante.

La mesure de *Rand* précédemment définie, compare le résultat de classification/segmentation avec une seule classification/segmentation de référence. Cependant, des versions plus génériques de la mesure de RAND ont été récemment proposées [Unnikrishnan 2007, Yang 2008] en prenant en considération plusieurs vérités-terrain.

La mesure de RAND probabiliste, notée  $\mathcal{PR}$  et proposée par UNNIKRISHNAN [Unnikrishnan 2005], est une nouvelle définition de la mesure de RAND qui compare le résultat de classification/segmentation automatique avec un ensemble d'images annotées manuellement par des experts (vérités-terrain). Cette métrique calcule une fonction de variabilité pour l'ensemble des classifications/segmentations de référence et attribue des poids non uniformes pour chaque paire de pixels.

Soient  $S$  l'image classée/segmentée automatiquement et  $\{VT_1, \dots, VT_Q\}$  l'ensemble de vérités-terrain. Considérons  $(L_i^S, L_j^S)$ , les labels de la paire de pixels  $(p_i, p_j)$  dans l'image  $S$  classée/segmentée automatiquement. Ce couple de labels est comparé aux couples

de labels correspondants dans chaque vérité-terrain  $(L_i^{VT_q}, L_j^{VT_q})$  (cf. équation 2.3).

$$\mathcal{PR}(S, VT_{1,...,Q}) = \frac{1}{\binom{n}{2}} \sum_{\substack{i,j \\ i \neq j}} [E(L_i^S = L_j^S).pb_{ij} + E(L_i^S \neq L_j^S).(1 - pb_{ij})] \quad (2.3)$$

où  $n$  est le nombre de pixels dans une image,  $E(L_i^S = L_j^S)$  est la probabilité d'avoir la paire de pixels  $(p_i, p_j)$  dans la même classe (donc ayant le même label) dans l'image  $S$ .  $pb_{ij}$  est la probabilité moyenne calculée sur toutes les images vérités-terrain (cf. équation 2.4).

$$pb_{ij} = \frac{1}{Q} \sum_{q=1}^Q E(L_i^{VT_q} = L_j^{VT_q}) \quad (2.4)$$

où  $E(L_i^{VT_q} = L_j^{VT_q})$  est la probabilité d'avoir la paire de pixels  $(p_i, p_j)$  dans la même classe (donc ayant le même label) dans l'image  $VT_q$ .

À l'inverse de la mesure  $\mathcal{R}$ , lorsque la valeur de  $\mathcal{PR}$  est proche de 1, la classification/segmentation obtenue est considérée comme précise, et lorsque la valeur de  $\mathcal{PR}$  est proche de 0, la classification/segmentation automatique est peu similaire à toutes les vérités-terrain considérées.

Une troisième définition de la mesure de *Rand*, notée  $\mathcal{NPR}$ , est introduite dans d'autres travaux de UNNIKRISHNAN [Unnikrishnan 2005, Unnikrishnan 2007] afin d'améliorer la mesure de RAND probabiliste. Cette mesure représente la normalisation de la mesure  $\mathcal{PR}$  par rapport à toutes les images de la base de test, en prenant en considération la variabilité des classifications/segmentations de référence. L'inconvénient majeur de cette métrique réside dans la complexité de calcul quand la taille de la base d'images est importante.

#### 2.2.2.2/ LA MESURE DE JACCARD

La mesure de JACCARD proposée par BEN-HUR [Ben-Hur 2002] mesure la similarité entre une image classée/segmentée  $S$  et une image vérité-terrain  $VT$  selon l'équation 2.5 donnée ci-dessous :

$$\mathcal{J}(S, VT) = 1 - \frac{n_{11}}{n_{11} + n_{01} + n_{10}} \quad (2.5)$$

Tout comme la mesure de RAND, cette métrique indique que la méthode de classification/segmentation est performante lorsque sa valeur est proche de zéro.

#### 2.2.2.3/ LA MESURE DE FOWLKES ET MALLOWS

La mesure de FOWLKES et MALLOWS a été déduite d'une mesure de similarité entre deux classifications hiérarchiques introduite dans les travaux [Fowlkes 1983].

$$\mathcal{F}(S, VT) = 1 - \sqrt{W_1(S, VT).W_2(S, VT)} \quad (2.6)$$

où les expressions  $W_1$  et  $W_2$  sont données respectivement par les équations 2.7 et 2.8 :

$$W_1(S, VT) = \sum_{i=1}^k \frac{n_{11}}{\frac{|r_i|(|r_i|-1)}{2}} \quad (2.7)$$

$$W_2(S, VT) = \sum_{j=1}^l \frac{n_{11}}{\frac{|r_j|(|r_j|-1)}{2}} \quad (2.8)$$

Tout comme les mesures de RAND et de JACCARD, cette métrique indique que la méthode de classification/segmentation est performante lorsque sa valeur est proche de zéro.

D'autres méthodes permettent de mesurer la similarité entre deux classifications/segmentations en quantifiant l'information partagée entre elles. Dans ce contexte, les deux métriques d'évaluation que nous allons introduire par la suite sont la mesure d'information mutuelle et la mesure de variation d'information.

#### 2.2.2.4/ LA MESURE D'information mutuelle

La mesure d'information mutuelle traduit la similarité entre deux segmentations/classifications en mesurant l'information partagée entre elles. Elle est calculée à l'aide d'outils issus de la théorie de l'information. Soient  $S$  la classification/segmentation automatique,  $VT$  une vérité-terrain et les deux vecteurs de variables aléatoires  $r_i \in S$  et  $r_j \in VT$ . La probabilité d'avoir un pixel  $p \in r_i$  (respectivement  $p \in r_j$ ) est donnée par la fonction de distribution marginale  $f(r_i)$  (respectivement  $f(r_j)$ ). La probabilité qu'un pixel  $p$  appartient en même temps à  $r_i$  et  $r_j$  est donnée par la fonction de distribution conjointe  $f(r_i, r_j)$ . L'information mutuelle entre la classification/segmentation automatique et la classification/segmentation de référence, connue aussi sous le nom de divergence de KULLBACK-LEIBLER [Kullback 1951] est donnée par l'équation 2.9 ci-dessous :

$$MI(S, VT) = \sum_{r_i \in S} \sum_{r_j \in VT} f(r_i, r_j) \log\left(\frac{f(r_i, r_j)}{f(r_i)f(r_j)}\right) \quad (2.9)$$

L'information mutuelle normalisée, introduite par [Strehl 2000], est donnée par l'équation 2.10 ci-dessous :

$$NMI(S, VT) = 1 - \frac{1}{\log(k \times l)} MI(S, VT) \quad (2.10)$$

où  $k$  et  $l$  sont les nombres de classes respectivement dans l'image classée/segmentée  $S$  et dans la vérité-terrain  $VT$ . Tout comme les métriques d'évaluation définies précédemment, la mesure  $NMI$  est bornée entre 0 et 1.

#### 2.2.2.5/ LA MESURE DE variation d'information

Une seconde forme de normalisation permet de générer une nouvelle métrique d'évaluation connue sous le nom de *variation d'information*. Cette métrique définie par

l'équation 2.11, a été introduite par MEILA [Meila 2003].

$$\mathcal{VI}(S, VT) = H(S) + H(VT) - 2MI(S, VT) \quad (2.11)$$

où  $H(S)$  est l'entropie de  $S$  et  $H(VT)$  est l'entropie de  $VT$  :

$$\begin{aligned} H(S) &= - \sum_{r_i \in S} f(r_i) \log f(r_i) \\ H(VT) &= - \sum_{r_j \in VT} f(r_j) \log f(r_j) \end{aligned} \quad (2.12)$$

Contrairement aux métriques déjà citées ci-dessus, les valeurs de la mesure  $\mathcal{VI}$  sont bornées à  $\log n$  où  $n$  est le nombre total de pixels dans l'image.

Par ailleurs, des indicateurs de performance peuvent être déduits en mesurant l'intersection des ensembles dans les classifications/segmentations à comparer. Dans ce contexte, VAN DONGEN et JIANG et al. ont proposé deux métriques différentes à savoir la mesure de VAN DONGEN (basée sur la distance de HAMMING) et la mesure *Bipartite Graph Matching* (basée sur l'appariement de graphes bipartis).

#### 2.2.2.6/ LA MESURE DE VAN DONGEN

Cette métrique notée  $\mathcal{D}$  doit son nom à VAN DONGEN [Dongen 2000] et est décrite par l'équation 2.13 ci dessous :

$$\mathcal{D}(S, VT) = 2n - D_H(S \Rightarrow VT) - D_H(VT \Rightarrow S) \quad (2.13)$$

où  $D_H(X \Rightarrow Y)$  désigne la distance de HAMMING [Hamming 1950] entre deux ensembles  $X$  et  $Y$  et est donnée par l'équation 2.14.

$$D_H(X \Rightarrow Y) = \sum_{x_i \in X} \sum_{y_j \in Y, y_j \neq x_i} |x_i \cap y_j| \quad (2.14)$$

Dans cette expression, la distance de HAMMING mesure la surface totale de chevauchement entre les régions qui appartiennent à la classification/segmentation automatique  $S$  et à la vérité-terrain  $VT$ .

#### 2.2.2.7/ LA MESURE DE *Bipartite Graph Matching*

JIANG et al. proposent cette nouvelle métrique basée sur l'appariement de graphes bipartis afin d'obtenir le meilleur appariement entre les classes/régions. Soient deux classifications/segmentations, la première notée  $S = \{r_i\}_{1 \leq i \leq k}$  et obtenue par un algorithme de classification/segmentation automatique et la deuxième notée  $VT = \{r_j\}_{1 \leq j \leq l}$  et annotée manuellement. Soit  $G = (V, E)$  un graphe biparti pondéré et non orienté qui peut être défini en ajoutant des arrêtes pondérées par les poids  $w_{ij} = |r_i \cap r_j|, \forall w_{ij} \neq 0$ . Les auteurs affirment que l'appariement maximal du graphe biparti pondéré est un sous-graphe de  $G$ , noté  $G'$ . Ce dernier est obtenu en supprimant les arrêtes par lesquelles les nœuds  $r_i$  et  $r_j$  possèdent au plus une seule arrête incidente et  $\sum w_{ij}$  est maximale. Compte tenu de cette définition, la métrique d'évaluation correspondante, notée  $\mathcal{BGM}$ , est donnée par l'équation 2.15 donnée ci-après :

$$\mathcal{BGM}(S, VT) = 1 - \frac{\sum w_{ij}}{n} \quad (2.15)$$

Maintenant que nous avons introduit les trois métriques basées sur les notions de classification (la mesure de RAND, la mesure de JACCARD et la mesure de FOWLKES et MALLOWS) et les deux métriques basées sur la théorie de l'information (les mesures d'*information mutuelle* et de *variation d'information*), nous allons décrire dans le reste de cette section deux autres métriques d'évaluation basées sur la perception humaine. Les auteurs adoptant cette stratégie d'évaluation considèrent généralement que l'erreur de segmentation est relative aux différents niveaux de détails de l'image segmentée. Par conséquent, toute métrique appartenant à cette catégorie est fortement sensible à la sur-segmentation ou à la sous-segmentation. Cette notion a son importance notamment pour les systèmes dont la complexité est une fonction du nombre de régions de l'image segmentée. En effet, une sur-segmentation doit être détectée puisque cela nécessite une étape de fusion de régions généralement coûteuse en temps de calcul. À l'inverse, une sous-segmentation de l'image doit aussi être détectée vu que les régions déterminées peuvent ne pas être fiables par rapport au contenu de l'image initiale. Par conséquent, une métrique d'évaluation doit refléter les besoins et les exigences demandées par l'application envisagée.

#### 2.2.2.8/ LA MESURE DE MARTIN

MARTIN [Martin 2002] propose deux métriques d'évaluation permettant d'examiner l'*uniformité locale* et l'*uniformité globale* entre deux segmentations différentes. Ces deux métriques sont l'*Erreur Locale de Cohérence* notée  $\mathcal{LCE}$  et l'*Erreur Globale de Cohérence* notée  $\mathcal{GCE}$ .

$$\mathcal{LCE}(S, VT) = \frac{1}{n} \sum_p \min\{E(p), E'(p)\} \quad (2.16)$$

$$\mathcal{GCE}(S, VT) = \frac{1}{n} \min\left\{\sum_p E(p), \sum_p E'(p)\right\} \quad (2.17)$$

avec

$$E(p) = \frac{|r_j \setminus r_i|}{|r_j|}$$

$$E'(p) = \frac{|r_i \setminus r_j|}{|r_i|}$$

où  $|\cdot|$  désigne le cardinal d'un ensemble quelconque. Soient  $X$  et  $Y$  deux régions données, l'expression  $X \setminus Y$  indique qu'un pixel  $p$  appartient à  $X$  mais pas à  $Y$ .  $n$  désigne le nombre total de pixels.

$\mathcal{LCE}$  et  $\mathcal{GCE}$  sont définies de telle sorte que lorsque nous comparons deux segmentations, plus les valeurs d'erreur sont faibles, voire nulles, plus les segmentations sont presque identiques.

### 2.2.2.9/ LA MESURE DE VINET

La mesure de VINET, notée  $\mathcal{V}$ , consiste à déterminer toutes les paires de régions qui ont un maximum de recouvrement et de définir la mesure comme le nombre de pixels qui ne participent pas au recouvrement [Cohen 1989]. En effet, cette métrique mesure l'asymétrie entre deux segmentations données. Soient  $S = \{r_i\}_{1 \leq i \leq k}$  la segmentation automatique et  $VT = \{r_j\}_{1 \leq j \leq l}$  la segmentation de référence. La première étape du calcul de la mesure de VINET consiste à générer une table de superposition obtenue à l'aide de l'équation 2.18.

$$t_{ij} = |r_i \cap r_j| \quad (2.18)$$

où  $|\cdot|$  désigne le cardinal d'un ensemble quelconque. La deuxième étape consiste à chercher les paires de régions qui ont un maximum de recouvrement dans la table définie par l'équation 2.18. La mesure de VINET est finalement donnée par l'équation 2.19 :

$$\mathcal{V} = 1 - \frac{1}{n} \times \sum_{s=1}^{\min(k,l)} C_s \quad (2.19)$$

avec

$$C_s = t_{i_s j_s} \geq t_{ij} \forall i \neq i_s, j \neq j_s$$

Les valeurs de cette métrique appartiennent à l'intervalle  $[0, 1]$ . Plus la valeur de  $\mathcal{V}$  est petite, meilleure est la segmentation.

Il existe d'autres métriques d'évaluation n'appartenant pas aux catégories précédemment détaillées (la catégorie basée sur les notions de classification, la catégorie basée sur la théorie de l'information et la catégorie basée sur la perception humaine). Nous décrivons ci dessous la mesure de YASNOFF et al. et la mesure de HUANG-DOM.

### 2.2.2.10/ LA MESURE DE YASNOFF

YASNOFF et al. considèrent que compter seulement le nombre de pixels mal segmentés est insuffisant pour mesurer la performance d'un algorithme de segmentation. Ils proposent une nouvelle mesure de performance tenant compte de la position du pixel (mal segmenté) [Yasnoff 1979]. Cette métrique utilise une distance entre le pixel mal segmenté dans la segmentation automatique  $S$  et la région à laquelle il appartient dans la segmentation de référence  $VT$ . L'expression de cette mesure, notée  $\mathcal{Y}$ , est donnée par l'équation 2.20.

$$\mathcal{Y} = \frac{100}{n} \times \sqrt{\sum_{p \in S^*} d^2(p)} \quad (2.20)$$

La sommation s'effectue sur l'ensemble  $S^*$  de pixels mal segmentés.  $d$  désigne la distance entre le pixel mal segmenté  $p$  et le pixel le plus proche de la région à laquelle il appartient dans l'image de référence.  $n$  est le nombre total de pixels dans l'image.

### 2.2.2.11/ LA MESURE DE HUANG-DOM

HUANG et DOM proposent dans leurs travaux [Huang 1995] une nouvelle métrique, notée  $\mathcal{HD}$ , qui consiste à mesurer la performance de l'algorithme de segmentation par rapport à une segmentation de référence en fonction du nombre, de la taille et de la localisation des régions dans l'image. Soient  $S = \{r_i\}_{1 \leq i \leq k}$  l'image segmentée de façon automatique et  $VT = \{r_j\}_{1 \leq j \leq l}$  la segmentation de référence, tel que  $r_i$  est l'ensemble des pixels appartenant à la région d'indice  $i$  de l'image segmentée  $S$ . Les auteurs associent à chaque ensemble  $r_i$ , un ensemble  $r_j$  tel que  $r_i \cap r_j$  est maximal. Nous rappelons que la distance de HAMMING  $D_H$  entre deux segmentations est donnée par l'équation 2.14 [Hamming 1950], ce qui revient à considérer la somme des surfaces d'intersection non maximales de toutes les régions. Par conséquent, la mesure  $\mathcal{HD}$  de HUANG-DOM est donnée par l'équation symétrique et normalisée :  $\mathcal{HD}$ .

$$\mathcal{HD}(S, VT) = 1 - \frac{D_H(VT \Rightarrow S) + D_H(S \Rightarrow VT)}{2n} \quad (2.21)$$

Pour retrouver l'indicateur de performance, il suffit de soustraire de 1 la valeur de la mesure  $\mathcal{HD}$ .

### 2.2.2.12/ AUTRES MÉTRIQUES D'ÉVALUATION

D'autres approches relatives à l'évaluation supervisée de la segmentation ont été citées dans la littérature. MEZARIS et al. [Mezaris 2003] ont proposé une approche dont l'idée est de fusionner, à l'aide d'une combinaison additive linéaire, deux mesures où la première mesure la performance des contours des régions et la deuxième s'intéresse au degré de sur-segmentation dans l'image segmentée. L'inconvénient majeur de cette technique réside dans le fait que la combinaison consiste à sommer simplement les deux mesures sans avoir recours à une pondération de cette somme.

CORREIA et PEREIRA décrivent dans leurs travaux [Correia 2003, Correia 2006] une méthode d'évaluation de segmentation d'images basée sur la similarité et la pertinence de la détection d'un objet. La similarité de l'objet est mesurée à l'aide d'un ensemble d'attributs : la fidélité de la forme (fonction du nombre de pixels mal segmentés), la similarité géométrique (la taille, la position, l'allongement et la compacité de l'objet), la similarité interne des contours (la différence moyenne des contours de SOBEL détectés au sein d'une région) et la similarité statistique (la différence de luminosité). La sélection ainsi que les méthodes de calcul de ces différentes mesures ne sont pas totalement justifiées. De plus, une pondération arbitraire a été utilisée pour combiner toutes ces mesures et obtenir la métrique finale de similarité d'objets.

USAMENTIAGA et al. [Usamentiaga 2006] proposent une métrique d'évaluation dont l'idée est un peu différente. Au lieu de comparer une paire de segmentations en mesurant le degré de chevauchement entre les régions, ils mesurent le degré de chevauchement entre les contours des régions. Parmi les inconvénients majeurs de cette métrique, la difficulté de détecter les contours des régions, la sensibilité aux fausses alarmes et les mauvaises détections.

La figure 2.7 et le tableau 2.1 illustrent les résultats d'une sélection de métriques d'évaluation (avec segmentation de référence) sur une image de la base de BERKELEY, en utilisant deux vérités-terrain issues de deux experts différents. Nous pouvons constater que le deuxième expert est plus exigeant que le premier puisqu'il définit plus de régions. Ceci justifie le fait que les évaluations sont majoritairement meilleures en considérant la première vérité-terrain. Nous constatons aussi que la variation dans les résultats avec les différentes vérités-terrain est relative à la métrique utilisée. Par exemple la mesure de MARTIN ( $LCE$  et  $GCE$ ) est moins sensible à la variation entre les segmentations de référence que la mesure d'information mutuelle  $MI$  et la mesure de VINET ( $\mathcal{V}$ ).



FIGURE 2.7 – Segmentation automatique par l'algorithme MeanShift et deux segmentations manuelles, d'une image de la base de BERKELEY.

Métriques	$\mathcal{R}$	$\mathcal{J}$	$\mathcal{F}$	$\mathcal{VI}$	$MI$	$BGM$	$LCE$	$GCE$	$\mathcal{V}$
Vérité terrain 1	0.2	0.51	0.34	1.98	0.83	0.66	0.17	0.23	0.31
Vérité terrain 2	0.47	0.76	0.59	2.88	0.36	0.75	0.1	0.17	0.78

TABLE 2.1 – Évaluation, avec différentes métriques, de la segmentation illustrée dans la figure 2.7.

Le lecteur intéressé par de plus amples informations sur les méthodes d'évaluation supervisée pourra consulter les références [Wirth 2006, Philipp-Foliguet 2006, Chau 2007] à partir desquelles nous avons extrait toutes les métriques détaillées dans cette section.

### 2.2.3/ MÉTHODES D'ÉVALUATION NON-SUPERVISÉE

Dans cette section, nous détaillons les approches d'évaluation non-supervisée. Contrairement aux approches d'évaluation de segmentation d'images avec segmentation de référence, les métriques d'évaluation non-supervisée mesurent la performance d'un algorithme de segmentation sans avoir recours à une image de vérité-terrain. En effet, le calcul des différentes métriques est basé sur l'hypothèse d'avoir des caractéristiques exploitables à partir de l'image segmentée automatiquement et qui peuvent refléter la qualité de la segmentation indépendamment d'une segmentation annotée par l'être humain. La justification de cette hypothèse s'appuie sur le fait que l'être humain lui-même est capable intuitivement de juger les performances d'un algorithme de segmentation en observant sa sortie sans avoir connaissance du contenu de l'image (la sur-segmentation est souvent facile à détecter par un observateur humain).

Les avantages de l'évaluation non-supervisée sont distincts de ceux de l'évaluation



supervisée et se manifestent par exemple dans le cas de la segmentation d'une image complexe. Le principal avantage réside dans la capacité de mesurer la performance de la segmentation sans avoir besoin d'une segmentation de référence. En effet, la création de vérités-terrain est coûteuse en temps et peut être pour certaines applications (médicale par exemple), une contrainte qu'il n'est pas toujours facile à lever. L'évaluation supervisée est subjective, fastidieuse, coûteuse, sans compter la difficulté de prendre en compte les vérités-terrain des différents experts.

L'hypothèse sur laquelle est fondée l'évaluation non-supervisée reste une approximation formelle et difficile à formaliser. Ceci justifie peut être le succès limité de cette catégorie de méthodes d'évaluation qui n'a pas connue un grand essor. Dans ce qui suit, nous allons rappeler les métriques d'évaluation non-supervisée les plus répandues.

### 2.2.3.1/ LA MESURE DE ZEBOUDJ

La mesure de ZEBOUDJ illustrée dans les travaux [Cocquerez 1995] consiste à évaluer la qualité d'une segmentation en mesurant deux critères de contraste : le *contraste intérieur* et le *contraste extérieur* aux régions mesurées sur un voisinage  $\mathcal{V}_p$  d'un pixel  $p$ . Le contraste entre le pixel  $p$  et un pixel voisin  $t \in \mathcal{V}_p$  est mesuré selon l'équation 2.22.

$$\text{Contraste}(p, t) = \frac{\|I(p) - I(t)\|}{L - 1} \quad (2.22)$$

où  $L$  est le maximum de l'image d'intensité  $I$  et  $\|\cdot\|$  désigne la valeur absolue.

Le *contraste intérieur* d'une région  $r_i$  dans une segmentation  $S = \{r_i\}_{1 \leq i \leq k}$  est donné par l'équation 2.23 ci dessous :

$$CI(r_i) = \frac{1}{|r_i|} \sum_{p \in r_i} \max\{\text{Contraste}(p, t), t \in \mathcal{V}_p \cap r_i\} \quad (2.23)$$

avec  $|\cdot|$  qui désigne le cardinal d'un ensemble donné.

Le *contraste extérieur* d'une région  $r_i$  dans la segmentation  $S$  est donné par l'équation 2.24 :

$$CE(r_i) = \frac{1}{lg_i} \sum_{p \in f_i} \max\{\text{Contraste}(p, t), t \in \mathcal{V}_p \text{ et } t \notin r_i\} \quad (2.24)$$

où  $f_i$  est la frontière de la région  $r_i$  et  $lg_i$  est la longueur de  $f_i$ . Le contraste d'une région est déduit de ses *contrastes extérieur* et *intérieur* et est donnée par l'équation 2.25 :

$$C(r_i) = \begin{cases} 1 - \frac{CI(r_i)}{CE(r_i)} & \text{si } 0 < CI(r_i) < CE(r_i) \\ CE(r_i) & \text{si } CI(r_i) = 0 \\ 0 & \text{sinon} \end{cases} \quad (2.25)$$

Par conséquent, le contraste global de l'image segmentée est donné par l'équation 2.26 :

$$\mathcal{Z} = \frac{1}{n} \sum_i |r_i| \cdot C(r_i) \quad (2.26)$$

où  $n$  est le nombre total de pixels dans l'image.

La valeur obtenue par la mesure  $\mathcal{Z}$  de ZEBOUJ est une valeur comprise dans l'intervalle  $[0, 1]$ . Plus le *contraste intérieur* est grand et plus le *contraste extérieur* est petit (la valeur de la mesure  $\mathcal{Z}$  est élevée), meilleure est la segmentation. L'inconvénient majeur de cette métrique d'évaluation est qu'elle n'est pas bien adaptée aux images très bruitées ou très texturées [Zhang 1996].

### 2.2.3.2/ LES MESURES DE LEVINE ET NAZIF

LEVINE et NAZIF définissent dans leurs travaux [Levine 1985] deux critères d'évaluation non-supervisée : le premier noté  $\mathcal{LN}_{intra}$  dédié à la mesure d'*uniformité intra-région* et est basé sur la somme des variances des régions. L'avantage de ce critère réside dans sa facilité de mise à jour si l'algorithme de segmentation possède des opérations de fusion ou de partition de régions. Ce critère est calculé à l'aide de l'équation 2.27 donnée ci-dessous :

$$\mathcal{LN}_{intra} = \sum_i \frac{\sigma_i^2}{\sigma_{max}} = \sum_i \sum_{p \in r_i} (I(p) - \frac{1}{|r_i|} \sum_{p \in r_i} I(p))^2 \quad (2.27)$$

avec

$$\sigma_{max} = \frac{(\max(I) - \min(I))^2}{2}$$

où  $r_i$  est une région de l'image segmentée  $S = \{r_i\}_{1 \leq i \leq k}$ ,  $I(p)$  désigne en général la valeur d'intensité ou la couleur au pixel  $p$ ,  $|r_i|$  est le cardinal (en pixel) de la région  $r_i$ ,  $\sigma_i$  est l'écart type de la région  $r_i$  et  $\sigma_{max}$  est un facteur de normalisation.

Le deuxième critère noté  $\mathcal{LN}_{inter}$ , dédié pour la mesure du *contraste inter-région*, consiste à calculer la somme des contrastes des régions  $r_i$  pondérées par leur aire comme le montre l'équation 2.28 :

$$\mathcal{LN}_{inter} = \frac{\sum_{r_i} |r_i| C_i}{\sum_{r_i} |r_i|} \quad (2.28)$$

avec

$$C_i = \sum_{r_{i'}} \frac{lg_{ii'}}{lg_i} c_{ii'}$$

$$c_{ii'} = \frac{\|m_i - m_{i'}\|}{\|m_i + m_{i'}\|}$$

où  $C_i$  représente le contraste de la région  $r_i$ ,  $r_{i'}$  est une région adjacente à  $r_i$ .  $c_{ii'}$  est le contraste entre deux régions adjacentes  $r_i$  et  $r_{i'}$ .  $m_i$  et  $m_{i'}$  sont les moyennes de l'image

d'intensité  $I$  dans les régions  $r_i$  et  $r_{i'}$ .  $lg_i$  est le périmètre de la région  $r_i$  et  $l_{ii'}$  est la longueur de la frontière entre les régions  $r_i$  et  $r_{i'}$ .

### 2.2.3.3/ LA MESURE DE DISSIMILARITÉ DE LIU ET YANG

La mesure de dissimilarité, notée  $\mathcal{LY}$  et proposée par LIU et YANG [Liu 1994], est proche dans son principe de calcul de la mesure de LEVINE et NAZIF [Levine 1985] avec une légère différence dans le calcul des écarts à la moyenne et dans la normalisation. Cette mesure caractérisée par ses faibles valeurs se calcule en fonction de trois paramètres : le nombre total de régions dans une segmentation  $S = \{r_i\}_{1 \leq i \leq k}$ , l'aire des régions et la couleur moyenne de chaque région. L'expression de la mesure  $\mathcal{LY}$  est donnée par l'équation 2.29 :

$$\mathcal{LY} = \frac{1}{1000 \times n} \sqrt{k} \sum_{i=1}^k \frac{e_i^2}{\sqrt{|r_i|}} \quad (2.29)$$

où  $n$  est le nombre total de pixels dans l'image,  $|r_i|$  est le cardinal de la région  $r_i$  donnant le nombre de pixels dans cette région et  $k$  est le nombre total de régions obtenues dans la segmentation  $S$ . La mesure  $e_i^2$  désigne la somme des distances euclidiennes entre les vecteurs couleurs des pixels de la régions  $r_i$  et le vecteur couleur attribué à la région  $r_i$  dans l'image segmentée (en général égal à la couleur moyenne des couleurs de la région).

L'inconvénient majeur de cette mesure est qu'elle pénalise les segmentations ayant un nombre de régions élevé (sur-segmentation) ou bien les segmentations ayant des régions non homogènes en couleur. Cet inconvénient est dû aux paramètres  $\sqrt{k}$  et  $\sqrt{|r_i|}$ .

### 2.2.3.4/ LA MESURE DE BORSOTTI

BORSOTTI et al. [Borsotti 1998] proposent une solution au problème de sensibilité de la mesure de LIU et YANG [Liu 1994] à la sur-segmentation en proposant une nouvelle mesure combinée notée  $\mathcal{B}$  et donnée par l'équation 2.30 :

$$\mathcal{B} = \frac{1}{1000 \times n} \sqrt{k} \sum_{i=1}^k \left[ \underbrace{\frac{e_i^2}{1 + \log |r_i|}}_{Terme_1} + \underbrace{\frac{N(r_i)^2}{|r_i|^2}}_{Terme_2} \right] \quad (2.30)$$

où  $n$ ,  $|r_i|$ ,  $k$ ,  $e_i^2$  désignent les mêmes paramètres déjà définis ci-dessus pour la mesure de LIU et YANG (cf. section 2.2.3.3).  $N(r_i)$  représente le nombre de régions ayant le même nombre de pixels que la région  $r_i$ .

Les améliorations apportées dans cette nouvelle métrique se manifestent dans la fusion de deux termes ( $Terme_1$  et  $Terme_2$ ) : le  $Terme_1$  pénalise les régions non homogènes (les régions deviennent non homogènes en général quand le nombre de régions dans l'image segmentée diminue) tandis que le  $Terme_2$  pénalise la sur-segmentation ayant beaucoup de régions de tailles similaires. Notons que plus la valeur de cette métrique est faible, meilleure est la segmentation.

### 2.2.3.5/ L'ERREUR QUADRATIQUE MOYENNE

L'erreur quadratique moyenne (MSE : Mean Squared Error), notée  $MSE$ , est une mesure utilisée pour calculer la performance d'une segmentation sans avoir recours à une segmentation de référence. En effet, cette métrique consiste à calculer comme son nom l'indique l'erreur quadratique moyenne entre une image initiale et une image restaurée ou améliorée. Une valeur faible du  $MSE$  signifie une faible erreur. L'expression du  $MSE$  est donnée par l'équation 2.31.

$$MSE = \frac{1}{n}(I - S)^2 \quad (2.31)$$

où  $n$  est le nombre total de pixels dans l'image initiale  $I$  et  $S$  est l'image segmentée.

L'inconvénient majeur de cette métrique réside dans sa sensibilité au bruit.

### 2.2.3.6/ AUTRES MÉTRIQUES D'ÉVALUATION NON-SUPERVISÉE

Dans la littérature, il existe d'autres métriques d'évaluation non-supervisée. Le critère de ZHANG et al. [Zhang 2004] proposent une mesure d'évaluation non-supervisée basée sur des outils de la théorie de l'information. Cette métrique utilise l'entropie comme un critère de désordre pour mesurer l'uniformité d'une région. CHEN et WANG définissent quant à eux deux mesures d'évaluation non-supervisée [Chen 2004] : une *erreur visuelle intra-région* qui mesure l'homogénéité moyenne de chaque région dans l'image et une *erreur visuelle inter-région* qui mesure l'écart-type entre les régions voisines dans l'image segmentée. La technique proposée par CHEN et WANG est proche dans son principe de la technique proposée par LEVINE et NAZIF [Levine 1985] (cf. section 2.2.3.2) avec des légères différences dans la pondération et la normalisation.

Le lecteur intéressé par de plus amples informations concernant les méthodes d'évaluation non-supervisée pourra consulter les références [Philipp-Foliguet 2006, Zhang 2008] à partir desquelles nous avons extrait toutes les métriques d'évaluation détaillées dans cette section.

Maintenant que nous avons rappelé les méthodes d'évaluation supervisée et non supervisée les plus couramment utilisées, nous allons présenter différentes bases d'images qui sont souvent utilisées pour tester et valider les méthodes de segmentation d'images. Deux bases d'images seulement ont été sélectionnées pour tester la performance de nos algorithmes, développés dans le cadre de nos travaux de thèse.

## 2.3/ LES BASES DE TESTS/APPRENTISSAGE/VALIDATION

Dans la littérature, plusieurs bases d'images ont été introduites par la communauté de la vision par ordinateur, ce qui a permis d'avoir une variété d'images couleur-texture, avec pour certaines des segmentations de référence établies par plusieurs experts. Nous allons dans la suite de ce chapitre, présenter les bases dédiées à la couleur et à la texture.

### 2.3.1/ BASES D'IMAGES COULEUR

Dans cette section, nous nous intéressons aux bases d'images couleur. Nous présentons trois bases d'images couleur fréquemment utilisées dans la littérature : la base d'images de MCGILL, la base d'images de BERKELEY et la base d'images de PASCAL VOC.

#### 2.3.1.1/ LA BASE D'IMAGES DE MCGILL

La base de tests de MCGILL [Olmos 2004], disponible en ligne à l'adresse <http://pirsquared.org/research/mcgilldb/welcome.html> fournit plus de 850 images naturelles couleur calibrées ayant des caractéristiques couleur-texture intéressantes. Cette base est divisée en neuf catégories (cf.figure 2.8) : fleurs, animaux, feuillages, textures, fruits, paysages, des images de l'hiver et d'autres avec la présence d'ombres. Des images illustratives extraites de quelques catégories de la base de MCGILL sont présentées sur la figure 2.8. Chaque catégorie contient entre 48 et 304 images de résolution  $786 \times 576$  pixels acquises par des caméras *Nikon Coolpix* numériques 5700. L'avantage de la base de MCGILL est qu'elle fournit des scènes de test complexes avec des images couleur non-homogènes et souvent très texturées pour la segmentation couleur-texture. Cependant, l'inconvénient majeur de cette base est qu'elle ne propose pas de segmentations de référence.

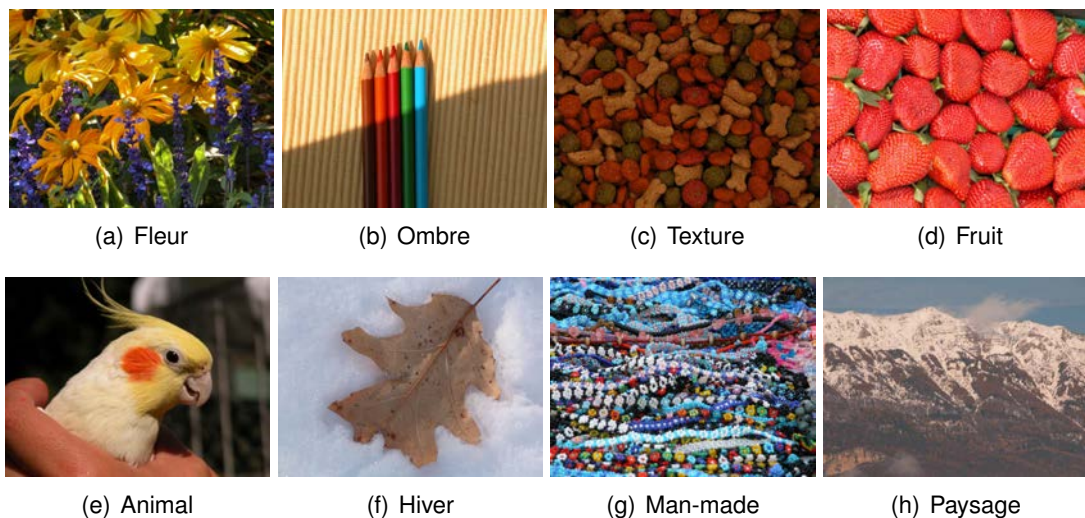


FIGURE 2.8 – Échantillons d'images issues de différentes catégories de la base de MCGILL.

#### 2.3.1.2/ LA BASE D'IMAGES DE BERKELEY

Dans le but de fournir une référence objective pour l'évaluation quantitative des algorithmes de segmentation, MARTIN et al. proposent la base d'images de BERKELEY (Berkeley Segmentation Dataset and Benchmark (BSDB)) [Berkeley , Martin 2001]. Cette base d'images contient initialement 300 images couleur, mais elle a été récemment élargie (BSDB500) pour contenir jusqu'à 500 images couleur [Arbelaez 2011]. La BSDB couvre une grande variété de scènes naturelles avec pour chacune d'elles un ensemble

de segmentations de référence annotées manuellement par différents experts. Les images de cette base ont une résolution de  $481 \times 321$  pixels et ont été sélectionnées de la base d'images de COREL (base d'images commerciale) en vérifiant qu'il existe dans chaque image au moins un objet visible. La figure 2.9 présente quelques images issues de la base de BERKELEY.



FIGURE 2.9 – Une sélection d'images avec quelques-unes de leurs segmentations de référence extraites de la base de BERKELEY.

La BSD500 est largement utilisée par la communauté de la vision par ordinateur et ceci pour deux raisons principales : la première est qu'elle dispose d'un grand nombre d'images couleur de scènes complexes. La deuxième raison est que pour chaque image de la base, plusieurs segmentations de références (entre 5 et 8) ont été réalisées par différents experts, ce qui facilite l'évaluation des algorithmes de segmentation. La base BSD500 [Arbelaez 2011], est disponible en ligne à l'adresse <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html>.



### 2.3.1.3/ LA BASE D'IMAGES DE PASCAL VOC

La base d'images de PASCAL VOC [Everingham 2009] est une base d'images contenant 14 743 images couleur avec leurs segmentations de référence. Cette base renferme des images de scènes réelles divisées en 20 catégories principales : avion, vélo, oiseau, bateau, cheval, voiture, chat, bus, chien, personne, etc. Chaque image de la base est munie d'un ensemble de vérités-terrain manuellement étiquetées. Cette base est disponible en ligne à l'adresse <http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2009/#devkit>. La figure 2.10 illustre quelques images et vérités-terrain de cette base.

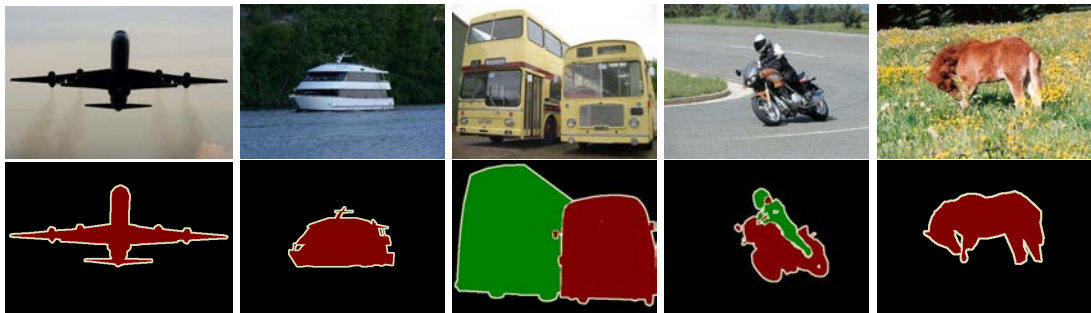


FIGURE 2.10 – Une sélection d'images de la base de PASCAL VOC.

### 2.3.2/ BASES D'IMAGES TEXTURE

Dans cette section, nous nous intéressons aux bases d'images dédiées pour la texture, à savoir la base d'images de VISTEX, la base d'images de OUTEX, la base d'images de PRAGUE et finalement la base d'images de librairie de texture.

#### 2.3.2.1/ LA BASE D'IMAGES DE VISTEX

VISION TEXTURE connue sous l'abréviation VISTEX [VisTex 1995] est une base d'images de texture souvent utilisée par la communauté de la vision par ordinateur. Elle a été proposée par l'Institut de Technologie de Massachusetts (MIT) en 1995. La figure 2.11 illustre quelques images extraites de cette base.

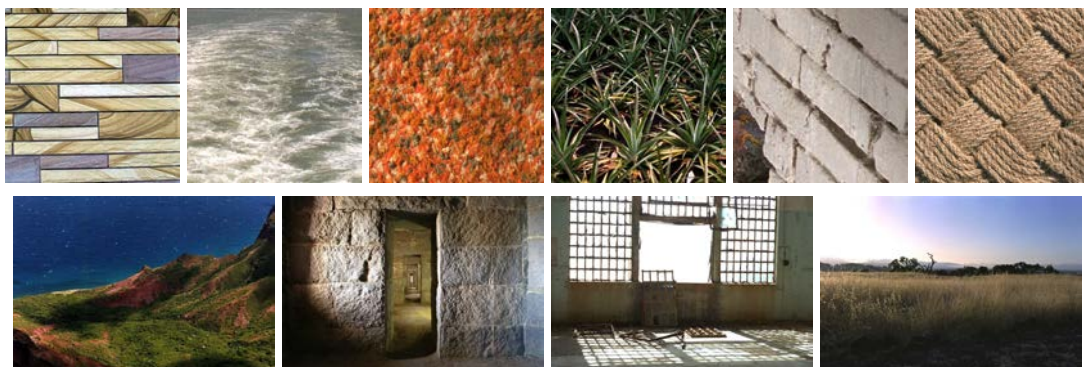


FIGURE 2.11 – Une sélection de textures uniformes et trois scènes naturelles extraites de la base de texture VISTEX.

VISTEX fournit plus de 100 échantillons de texture couleur de résolution  $128 \times 128$  pixels et  $512 \times 512$  pixels ainsi que des scènes naturelles complexes. Les textures de la base VISTEX appartiennent à plusieurs prototypes (herbes, eaux, toiles, briques, bâtiments et nuages) et ont été souvent exploitées pour construire des images synthétiques de textures composées, qui à leur tour ont été utilisées pour mesurer la performance des algorithmes de segmentation couleur/texture. La base de VISTEX est disponible en ligne à l'adresse <http://vismod.media.mit.edu/pub/VisTex/VisTex.tar.gz>.

### 2.3.2.2/ LA BASE D'IMAGES DE OUTEX

La base de données de texture OUTEX [Ojala 2002] a également eu du succès auprès des chercheurs en vision par ordinateur. En effet, cette base offre une grande collection de textures couleur (toile, tapis, bois, sable, tuiles, etc) qui peuvent être utilisées à la fois pour l'évaluation de la classification de texture et les algorithmes de segmentation. Cette base d'images de texture est disponible dans le site <http://www.outex.oulu.fi/>. La figure 2.12 illustre quelques images de cette base.

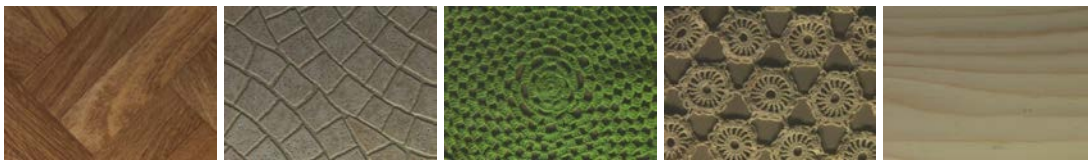


FIGURE 2.12 – Une sélection d'images de la base de texture OUTEX.

### 2.3.2.3/ LA BASE D'IMAGES DE PRAGUE

HAINDL et MIKES ont proposé en 2008 [Haindl 2008], une base d'images connue sous le nom de *Prague Texture Segmentation Data Generator and Benchmark*. Elle fournit plus de 1000 mosaïques texturées. On y retrouve six catégories de mosaïques : mono-spectrale, multi-spectrale, fonction de texture bidirectionnelle (Bidirectional Texture Function (BTF)), invariance de rotation, invariance d'échelle et invariance de luminosité. Cette base d'images est librement accessible à l'adresse <http://mosaic.utia.cas.cz/>. La figure 2.13 illustre quelques images extraites de cette base.



FIGURE 2.13 – Une sélection de textures mono-spectrales et multi-spectrales de la base de PRAGUE.

Les auteurs précisent que les segmentations de référence réalisées manuellement sont généralement caractérisées par la subjectivité humaine. Par conséquent, ils construisent des scènes synthétiques en mosaïque pour que leurs segmentations de référence ne



soient pas ambiguës. Notons aussi que dans le but d'aider l'utilisateur dans le processus de génération de scènes synthétiques, les auteurs ont développé un algorithme basé sur un générateur aléatoire polygone de VORONOI [Fortune 1987].

#### 2.3.2.4/ LA BASE D'IMAGES DE LA LIBRAIRIE DE TEXTURE

Accessible à l'adresse <http://textures.forrest.cz>, la librairie de texture est une collection de textures couleur qui a été mise à jour en 2009. Elle est constituée de 17 différentes catégories de textures couleur avec plus de 153 images dans chaque catégorie. La figure 2.14 illustre quelques images



FIGURE 2.14 – Différentes textures extraites de la librairie de texture.

Pour plus amples informations sur les bases d'images couleur et texture qui ont été utilisées dans la littérature pour l'évaluation et la validation des algorithmes de segmentation, ILEA et WHELAN réalisent une étude plus détaillée dans leurs travaux [Ilea 2011] que le lecteur intéressé pourra consulter.

Maintenant que nous avons introduit les bases d'images que nous pouvons utiliser pour valider nos algorithmes, nous allons, dans ce qui suit, réaliser une étude comparative des performances des méthodes de segmentation déjà introduites dans le chapitre 1 via une sélection de métriques d'évaluation présentées dans ce chapitre en considérant deux bases d'images sélectionnées.

## 2.4/ COMPARATIF DES MÉTHODES DE SEGMENTATION COULEUR/TEXTURE

Dans cette section, nous évaluons et comparons la performance d'un ensemble de méthodes de segmentation précédemment décrites (cf. Chapitre 1) à l'aide d'un ensemble de métriques d'évaluation supervisée et non supervisée. Les méthodes à comparer ont été testées seulement sur deux bases d'images choisies parmi celles que nous avons présentées dans la section précédente. Avant de détailler l'étude comparative, nous justifions les choix des bases de tests et des métriques d'évaluation.

### 2.4.1/ CHOIX DE BASES DE TESTS

Afin de comparer les différentes méthodes de segmentation décrites dans le chapitre 1, nous avons choisi d'effectuer les tests sur deux bases d'images. La première est la base d'images couleur de BERKELEY [Arbelaez 2011]. Nous avons choisi cette base

d'images couleur puisqu'elle est la plus citée dans la littérature et est fréquemment utilisée pour l'évaluation des algorithmes de segmentation [Martin 2004, Chen 2005, Wang 2006, Ilea 2008, Yang 2008, Yanga 2008, Kim 2009, Krinidis 2009, Hedjam 2009, Khan 2009, Mignotte 2010]. Elle présente aussi l'avantage de fournir 500 scènes naturelles complexes intéressantes pour l'évaluation. La deuxième base d'images que nous avons choisie, est la base de textures couleur VISTEX [VisTex 1995]. Tout comme la base de BERKELEY, nous avons choisi cette base pour sa fréquente utilisation dans la littérature [Mirmehdi 2000, Chen 2002, Nammalwar 2004, Freixenet 2004, Wang 2006, Kato 2006, Ilea 2008, Kim 2009]. Nous avons sélectionné seulement 33 échantillons de textures couleur de cette base de test. Nous considérons comme de nombreux auteurs, que ce nombre d'images est acceptable pour pouvoir évaluer la performance d'une méthode de segmentation. En effet, le nombre d'images de la base de VISTEX généralement utilisé varie entre 5 et 46 images (mosaïques et naturelles). Nous citons à titre d'exemple, MIRHMEDI et PETROU [Mirmehdi 2000] qui ont utilisé 27 mosaïques et 4 images naturelles, NAMMALWAR et al. [Nammalwar 2004] qui ont utilisé 46 images de textures entre mosaïques et images naturelles, FREIXENET et al. [Freixenet 2004] qui ont utilisé 9 mosaïques, Kim et Hong [Kim 2009] qui ont utilisé seulement 6 mosaïques et ILEA et WHELAN [Ilea 2008] qui ont utilisé 33 mosaïques pour évaluer leur méthode CTex.

## 2.4.2/ CHOIX ET DÉVELOPPEMENT DES MÉTRIQUES D'ÉVALUATION

Pour évaluer les résultats de segmentation, nous avons choisi un ensemble de métriques d'évaluation supervisée telles que la mesure de RAND probabiliste  $PR$ , la mesure de JACCARD  $\mathcal{J}$ , la mesure de FOWLKES et MALLOWS  $\mathcal{F}$ , la mesure de *Bipartite Graph Matching*  $BGM$ , la mesure de variation d'information  $VI$ , la mesure d'information mutuelle normalisée  $NMI$ , les mesures de MARTIN  $LCE$  et  $GCE$  et la mesure de VINET. Ont aussi été utilisées, des métriques d'évaluation non supervisée telles que la mesure de  $MSE$ , la mesure de ZBOUDJ  $\mathcal{Z}$  et la mesure de BSORBOTTI  $\mathcal{B}$ . Le critère de sélection des métriques d'évaluation est basé sur leur occurrence d'utilisation dans la littérature et leur capacité à mesurer la performance d'un algorithme de segmentation.

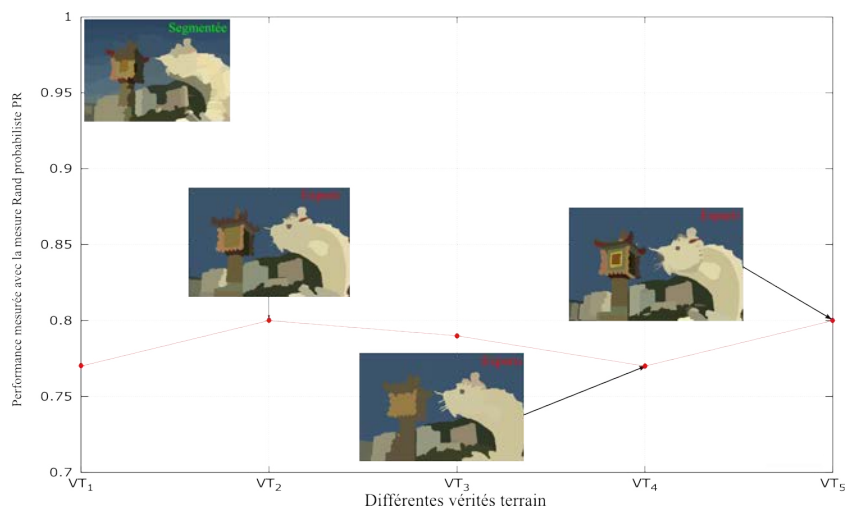


FIGURE 2.15 – Exemple de variation de la performance de la segmentation sur une image en fonction de cinq vérités-terrain.

Dans le cadre de nos tests, nous avons choisi d'appliquer un large nombre de métriques d'évaluation pour deux raisons. La première est d'acquérir une meilleure compréhension de ces métriques existantes (supervisées et non supervisées). Nous tenons à développer une compréhension de la portée de ces mesures, en particulier, des valeurs qui indiquent une segmentation précise, et des valeurs qui indiquent une segmentation inexacte. La deuxième raison est que nous tenons également à évaluer la capacité d'un algorithme de segmentation d'images de manière la plus objective possible, en utilisant l'ensemble des métriques sans privilège.

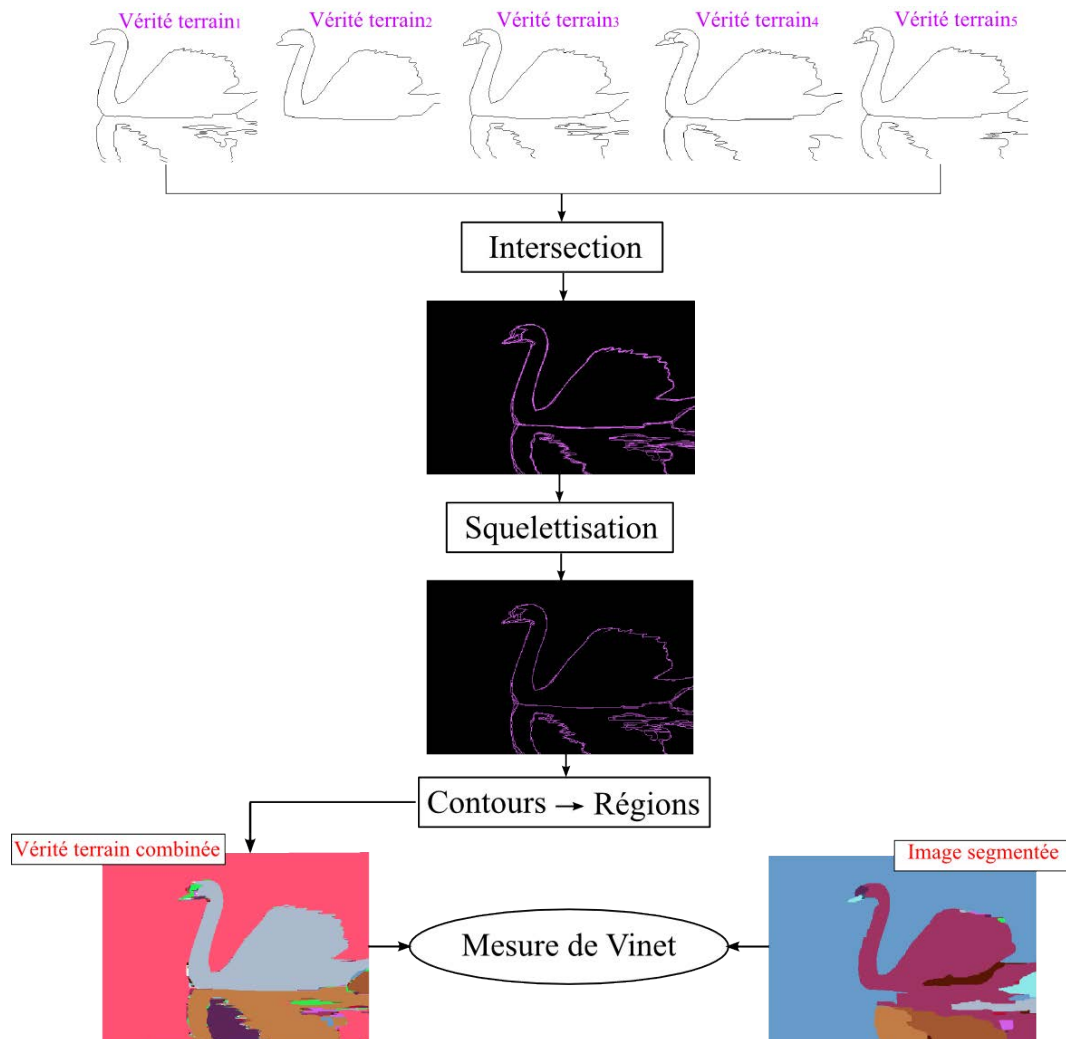


FIGURE 2.16 – Processus de l'évaluation par les mesures de VINET.

Pour les mesures  $\mathcal{J}$ ,  $\mathcal{F}$ ,  $\mathcal{BGM}$ ,  $\mathcal{NMI}$ , et  $\mathcal{V}$ , nous avons choisi d'utiliser la formule de performance au lieu de la mesure de dissimilarité. Par conséquent, plus ces métriques sont proches de 1, meilleure est la segmentation. Les mesures  $\mathcal{PR}$ ,  $\mathcal{J}$ ,  $\mathcal{F}$ ,  $\mathcal{BGM}$ ,  $\mathcal{VI}$  et  $\mathcal{NMI}$  ont été calculées pour chaque vérité-terrain fournie dans la base de BERKELEY. La figure 2.15 illustre à titre d'exemple, la variation d'évaluation d'une segmentation effectuée par l'algorithme JSEG sur une image de la base de BERKELEY en fonction de chaque expert (ie. de chaque segmentation de référence). L'évaluation a été réalisée avec la mesure de RAND probabiliste  $\mathcal{PR}$ .

Contrairement aux autres mesures d'évaluation supervisée qui prennent en entrée l'image segmentée et un ensemble de segmentations de référence, la mesure de VINET  $\mathcal{V}$  prend quant à elle deux images : l'image originale et une seule segmentation de référence. Cependant, à chaque image de la base de BERKELEY, est associé un ensemble de segmentation de référence. Afin de prendre en considération toutes les segmentations de référence, mais sans privilégier aucune d'entre elles, nous avons fait le choix d'effectuer une intersection des images de contours de ces vérités-terrain. Nous avons par la suite transformé l'image de contours résultante (l'intersection des contours des vérités-terrain) en une image de régions. La figure 2.16 illustre le processus d'évaluation avec la mesure de VINET. Notons qu'une étape de squelettisation a été ajoutée afin d'affiner les contours qui deviennent épais à l'issue de l'opération d'intersection.

### 2.4.3/ EXPÉRIMENTATION ET DISCUSSIONS

Nous rappelons que dans cette section, nous nous intéressons à l'évaluation de quelques-unes des méthodes de segmentation basées sur la combinaison des informations couleur/texture. La segmentation basée sur la combinaison couleur/texture fixe [Angulo 2007] dépend des gradients combinés et du facteur de pondération ( $\alpha$ ). Par conséquent, cette approche a été évaluée en fonction de ces deux paramètres. Nous rappelons rapidement les couples gradient couleur/texture testés et décrits plus en détails dans le chapitre 3. Le tableau 2.2 résume les notations de chaque combinaison couleur/texture testée, afin de les retrouver plus aisément dans les tableaux de résultats.

Gradient texture	Gradient couleur	
	Morphologique	Di-Zenzo
filtre ASF	AM	AZ
filtre de Gabor	GM	GZ
1 <sup>er</sup> paramètre d'HARALICK	H1M	H1Z
2 <sup>ème</sup> paramètre d'HARALICK	H2M	H2Z
5 <sup>ème</sup> paramètre d'HARALICK	H5M	H5Z

TABLE 2.2 – Désignation des différentes combinaisons couleur/texture testées.

	Évaluation	$\alpha^*$
	AM	
$\mathcal{PR}$	<b>0.8</b>	0.2, 0.5, 0.7, 1
$\mathcal{J}$	0.23	0, 0.2
$\mathcal{F}$	0.40	0.2
$\mathcal{BGM}$	0.15	0, 0.2
$\mathcal{NMI}$	<b>0.13</b>	0.5, 0.7, 1
$\mathcal{VI}$	4.44	0.2
$\mathcal{LCE}$	0.22	0.5, 0.7, 1
$\mathcal{GCE}$	0.31	0.7
$\mathcal{V}$	<b>0.96</b>	0, 0.2
$\mathcal{MSE}$	320.52	0.7
$\mathcal{Z}$	<b>0.48</b>	1
$\mathcal{B}$	0.10	0.5, 0.7, 1

	Évaluation	$\alpha^*$	Évaluation	$\alpha^*$	Évaluation	$\alpha^*$
	AZ		GM		GZ	
$PR$	<b>0.8</b>	0.2, 0.5	<b>0.8</b>	0.2, 0.5, 0.7, 1	0.79	0, 0.2, 0.5, 0.7, 1
$J$	0.23	0	0.20	0	0.20	0
$F$	0.39	0	0.36	0, 1	0.36	0
$BGM$	0.15	0	0.14	0	0.14	0
$NMI$	<b>0.13</b>	0.5, 0.7, 1	<b>0.13</b>	0.5, 0.7, 1	<b>0.13</b>	0.5, 0.7, 1
$VI$	4.51	0	4.60	1	4.61	0
$LCE$	0.22	0.5, 0.7	0.22	0.7, 1	0.23	0.5, 0.7
$GCE$	0.30	0.7	0.31	0.7	0.30	0.7
$V$	<b>0.96</b>	0	<b>0.96</b>	0, 0.2	<b>0.96</b>	0, 0.2
$MSE$	311.84	0.7	318.84	0.7	309.76	0.7
$Z$	0.39	0.5, 0.7, 1	0.48	1	0.39	1
$B$	0.13	0, 0.2, 0.5, 0.7	0.10	0.7, 1	0.12	0, 0.5, 0.7
	H1M		H1Z		H2M	
$PR$	<b>0.8</b>	1	0.79	1	<b>0.8</b>	0.5, 0.7, 1
$J$	0.19	1	0.13	1	<b>0.29</b>	0.2
$F$	0.36	1	0.28	1	<b>0.45</b>	0, 0.2
$BGM$	0.13	1	0.10	1	<b>0.24</b>	0
$NMI$	<b>0.13</b>	0, 0.7, 1	<b>0.13</b>	0, 0.7, 1	<b>0.13</b>	0.7, 1
$VI$	4.6	1	5.20	1	<b>3.96</b>	0
$LCE$	0.22	1	0.24	0.7, 1	<b>0.18</b>	0
$GCE$	0.32	1	0.33	1	<b>0.29</b>	0
$V$	<b>0.96</b>	0, 0.2, 0.5, 0.7	<b>0.96</b>	0, 0.2, 0.5, 0.7	<b>0.96</b>	0, 0.2
$MSE$	328.18	1	348.31	0.7	<b>328.18</b>	1
$Z$	<b>0.48</b>	1	0.39	1	<b>0.48</b>	1
$B$	0.10	1	0.14	1	<b>0.10</b>	0.2, 0.5, 0.7, 1
	H2Z		H5M		H5Z	
$PR$	<b>0.8</b>	0.2, 0.5	<b>0.8</b>	0.5, 0.7, 1	<b>0.8</b>	0.2, 0.5
$J$	0.28	0	0.22	0	0.22	0
$F$	<b>0.45</b>	0	0.38	0	0.38	0
$BGM$	<b>0.24</b>	0	0.16	0	0.16	0
$NMI$	<b>0.13</b>	0.5, 0.7, 1	<b>0.13</b>	0.5, 0.7, 1	<b>0.13</b>	0.2, 0.5, 0.7, 1
$VI$	<b>3.96</b>	0	4.43	0	4.43	0
$LCE$	<b>0.18</b>	0	0.22	0, 0.2, 0.5, 0.7, 1	0.22	0, 0.2, 0.5, 0.7, 1
$GCE$	<b>0.29</b>	0	0.32	0.7, 1	0.31	0.5, 0.7
$V$	<b>0.96</b>	0	<b>0.96</b>	0	<b>0.96</b>	0
$MSE$	327.19	0.7	323.79	0.7	<b>297.68</b>	0.7
$Z$	0.43	0.5	<b>0.48</b>	1	0.40	0.7
$B$	0.11	0	<b>0.09</b>	0.2, 0.5	0.10	0, 0.2

TABLE 2.3 – Performances moyennes (calculées sur 200 images de la base de BERKELEY avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture) de la segmentation par combinaison couleur/texture fixe [Angulo 2007] ( $\alpha^*$  correspond à la valeur optimale du paramètre de combinaison  $\alpha$ ).



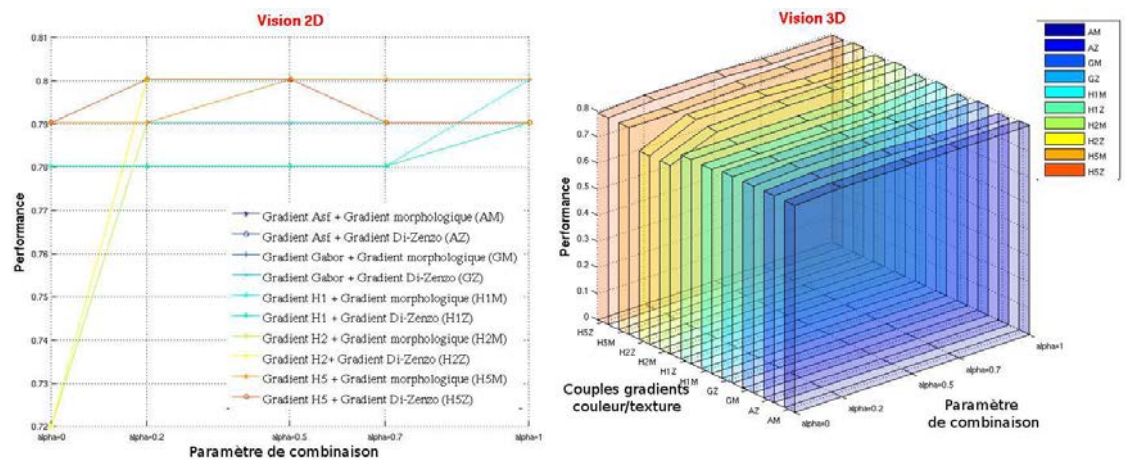
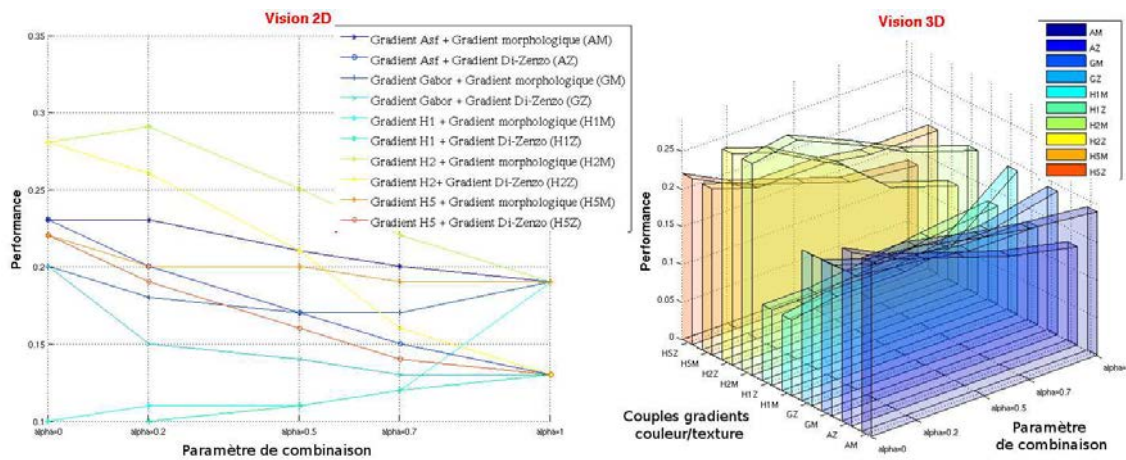
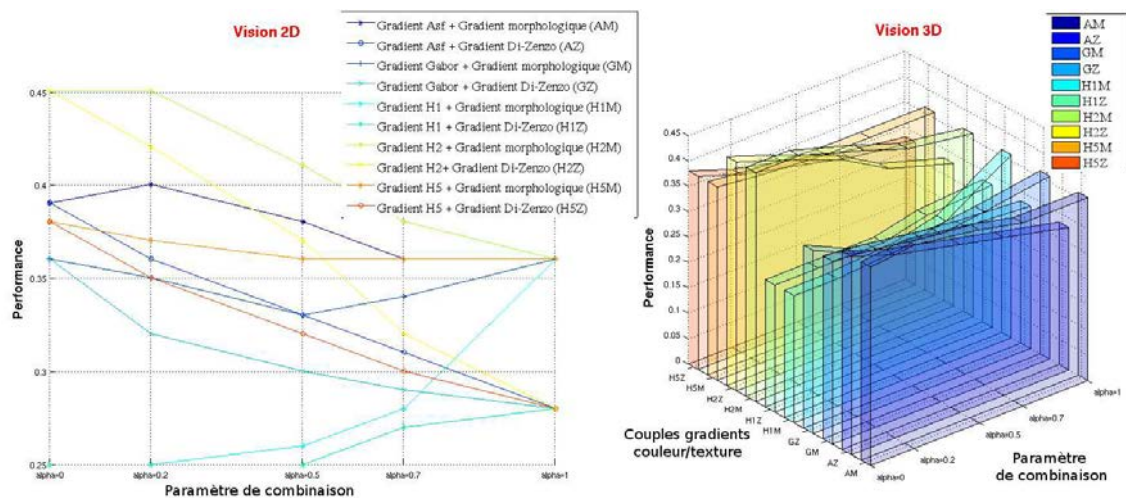
(a) Mesure de RAND probabiliste ( $\mathcal{PR}$ )(b) Mesure de JACCARD ( $\mathcal{J}$ )(c) Mesure de FOWLKES et MALLOWS ( $\mathcal{F}$ )

FIGURE 2.17 – Évaluation moyenne (avec les métriques de RAND  $\mathcal{PR}$ , de JACCARD  $\mathcal{J}$  et de FOWLKES et MALLOWS  $\mathcal{F}$  sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison  $\alpha$ .

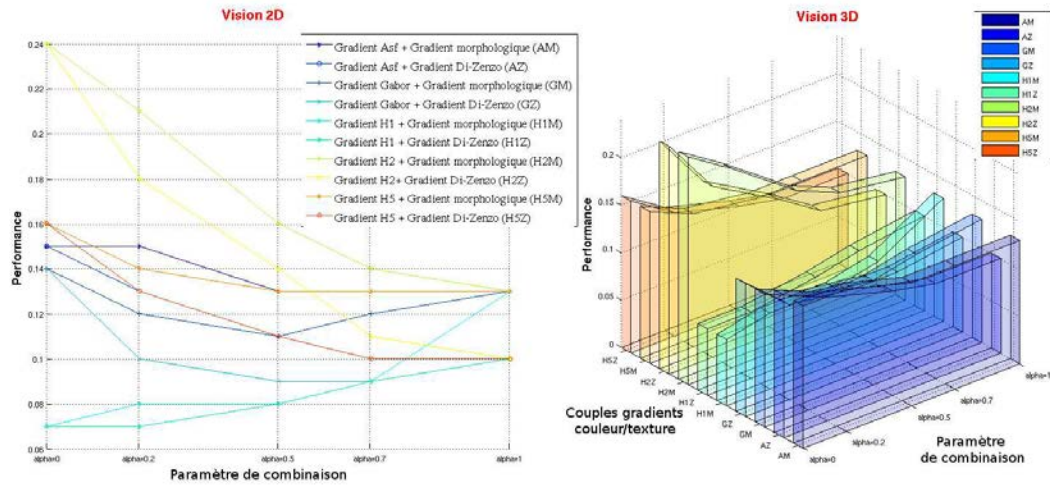
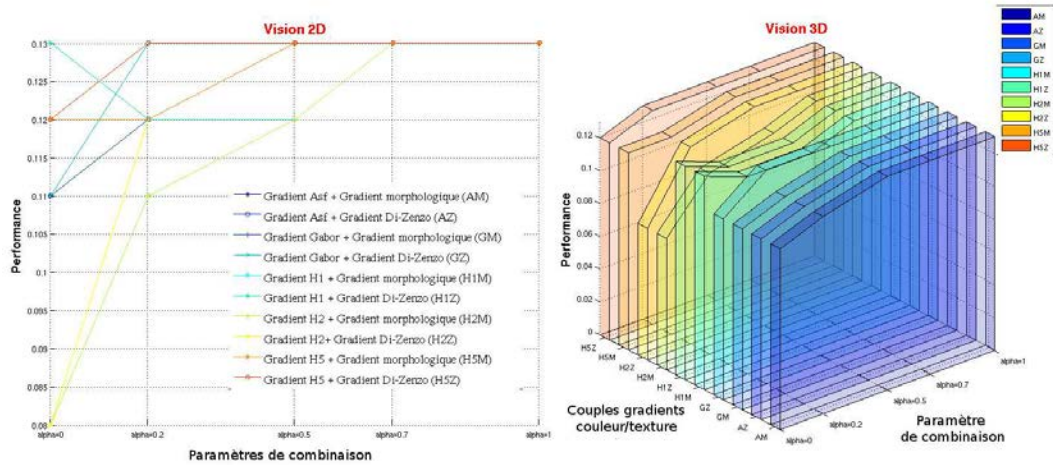
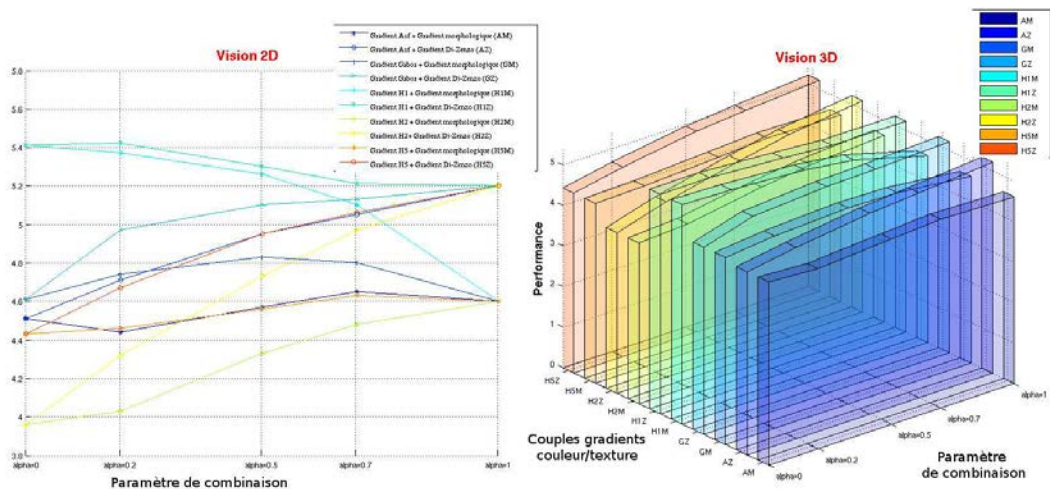
(a) Mesure de *Bipartite Graph Matching* ( $BGM$ )(b) Mesure d'Information mutuelle normalisée ( $NMI$ )(c) Mesure de *Variation d'information* ( $VI$ )

FIGURE 2.18 – Évaluation moyenne (avec les métriques de Bipartite Graph Matching  $BGM$ , d'information mutuelle normalisée  $NMI$  et de variation d'information  $VI$  sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison  $\alpha$ .



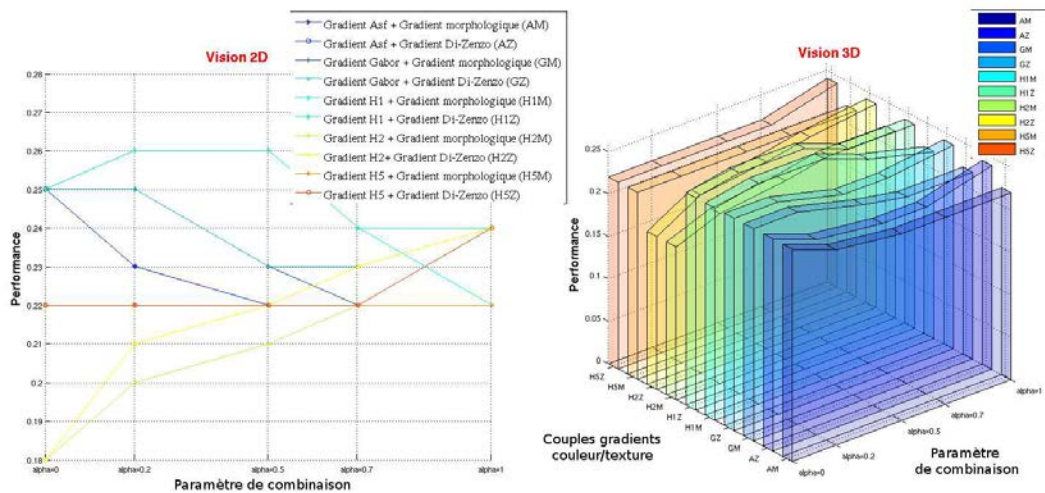
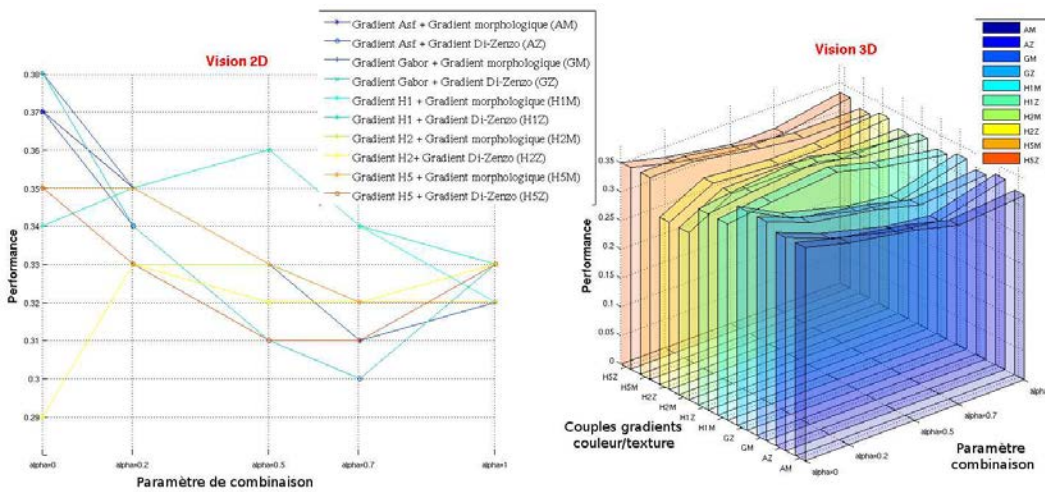
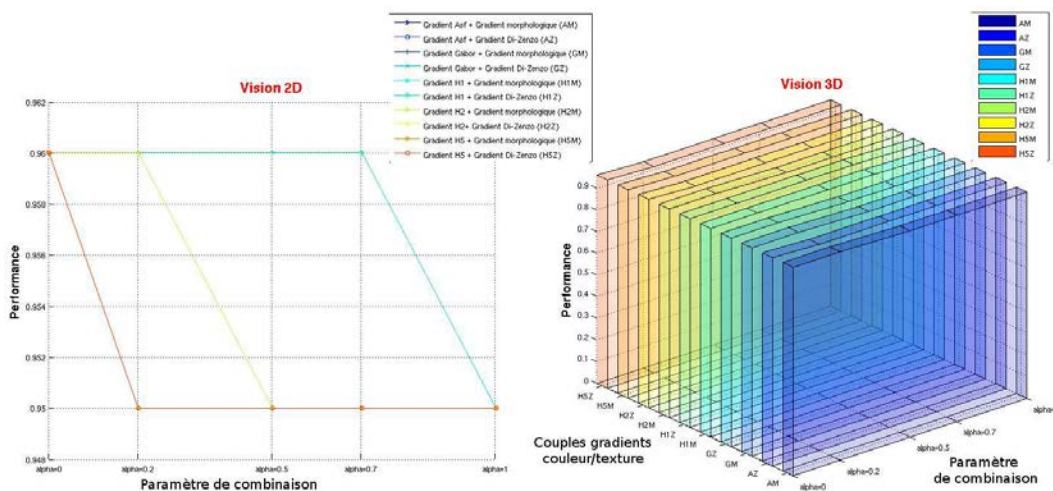
(a) Mesure de MARTIN  $\mathcal{LCE}$ (b) Mesure de MARTIN  $\mathcal{GCE}$ (c) Mesure de VINET ( $\mathcal{V}$ )

FIGURE 2.19 – Évaluation moyenne (avec les métriques de MARTIN  $\mathcal{LCE}$ ,  $\mathcal{GCE}$  et de VINET  $\mathcal{V}$  sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texte en fonction du paramètre de combinaison  $\alpha$ .



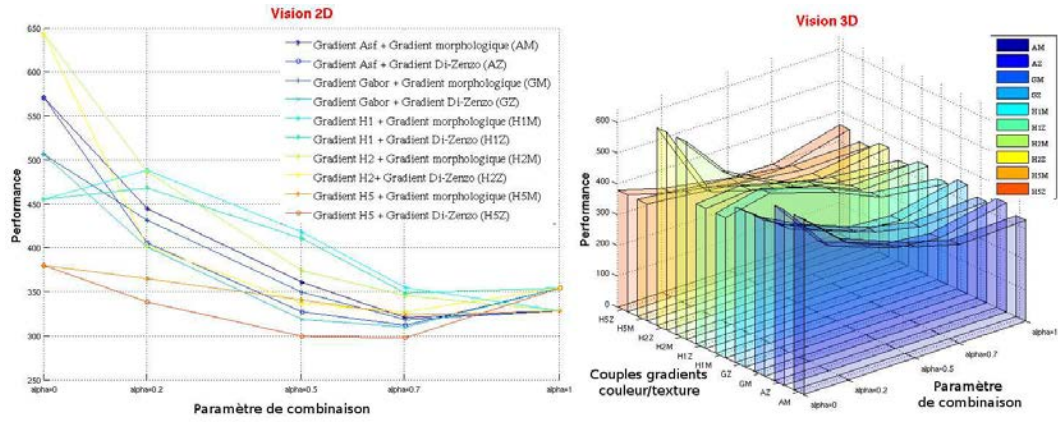
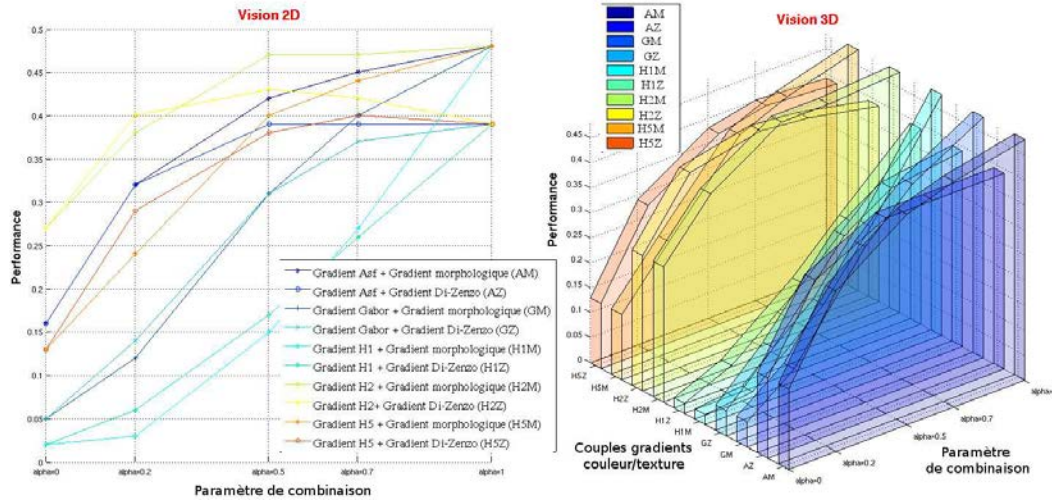
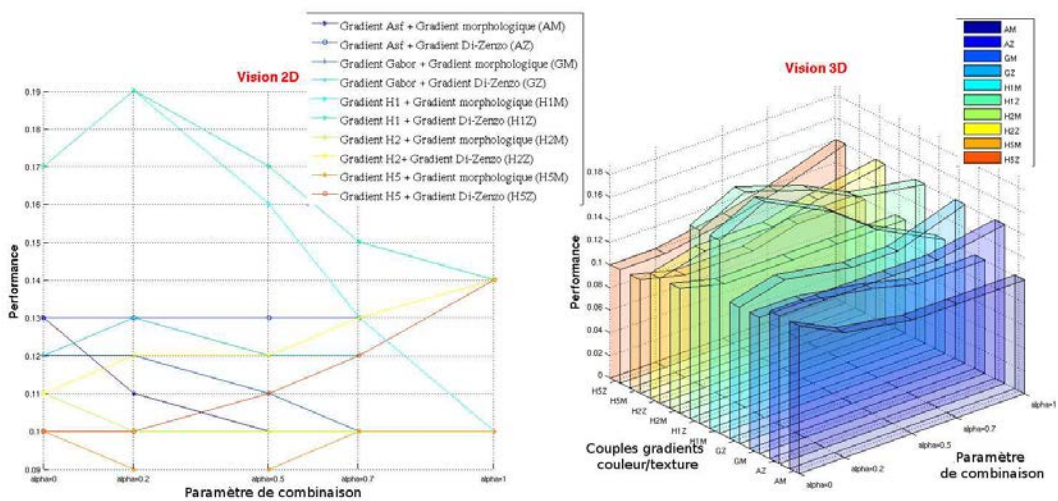
(a) Mesure d'Erreur quadratique moyenne ( $MSE$ )(b) Mesure de ZEBOUJ ( $Z$ )(c) Mesure de BORSOTTI ( $B$ )

FIGURE 2.20 – Évaluation moyenne (avec les métriques d'Erreur quadratique moyenne  $MSE$ , de ZEBOUJ  $Z$  et de BORSOTTI  $B$  sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison  $\alpha$ .

Le tableau 2.3 et les figures 2.17, 2.18, 2.19 et 2.20 illustrent la performance de la méthode de segmentation basée sur une combinaison fixe de gradients couleur/texture. Plus précisément, les figures 2.17, 2.18, 2.19 et 2.20 présentent sous deux visions (2D et 3D), les évaluations moyennes calculées sur 200 images de la base de BERKELEY (BSD500) avec les métriques considérées. La performance de la segmentation est présentée séparément pour chaque métrique en fonction du facteur de pondération  $\alpha$  et des couples gradient couleur/gradient texture testés. Nous retenons dans le tableau 2.3, pour chaque métrique d'évaluation utilisée, aussi bien les valeurs du facteur de pondération ainsi que les combinaisons couleur/textures optimales correspondant aux meilleures performances moyennes de la méthode de segmentation par combinaison fixe de gradients couleur/textures.











Image initiale	Combinaison fixe H5Z					Image segmentée avec $\alpha^*$
	$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 1$	
	<b>189.7</b>	198.6	199.4	204.4	190.5	
	169.3	159.8	<b>114.4</b>	127.1	116.4	
	244.1	138.3	131.6	<b>117.7</b>	135.1	
	123.2	115.4	112.6	79.1	<b>78.7</b>	
	75.2	68.1	<b>65.8</b>	66.6	73.4	

TABLE 2.4 – Résultats (avec une évaluation par  $MSE$ ) de la segmentation par combinaison fixe couleur/textures sur cinq images de la base de BERKELEY ( $\alpha^*$  correspond à la valeur optimale du paramètre de combinaison  $\alpha$ ).

D'après le tableau 2.3, nous pouvons constater que la plupart des métriques d'évaluation testées favorisent la combinaison H2M (correspondant au couple : gradient texture basé sur le 2<sup>ème</sup> paramètre d'HARALICK et gradient morphologique couleur). La deuxième constatation est que nous avons une valeur différente du paramètre de combinaison  $\alpha$  pour chaque métrique utilisée. Par exemple, la mesure  $\mathcal{J}$  favorise la combinaison H2M avec un paramètre  $\alpha = 0.2$  tandis que l'évaluation avec  $\mathcal{VI}$  et l'évaluation avec  $\mathcal{Z}$  sont optimales respectivement avec  $\alpha = 0$  et  $\alpha = 1$ . Nous pouvons donc conclure en mettant en évidence la difficulté de choisir une valeur adaptée pour le paramètre de combinaison  $\alpha$ .

Pour une meilleure visualisation, le tableau 2.4 présente les résultats de segmentation basée sur une combinaison fixe couleur/textures H5Z, évaluée avec la métrique  $MSE$  pour différentes valeurs de  $\alpha$ . Nous constatons que  $\alpha$  doit être redéfini pour chaque image afin d'assurer la meilleure segmentation possible.

Tout comme la précédente approche, les segmentations d'images par combinaison adaptative et par combinaison supremum des informations de couleur et de texture proposées dans les travaux de COHEN et al. [Cohen 2009], dépendent du même paramètre de pondération, mais qui est défini de manière adaptative en fonction du contenu de l'image. Nous abordons alors les tests avec la même stratégie d'évaluation que précédemment, en présentant pour chaque métrique utilisée, la qualité de la segmentation obtenue en fonction de la combinaison adaptative/supremum de gradients couleur/texture utilisés.

Combinaison adaptative										
	AM	AZ	GM	GZ	H1M	H1Z	H2M	H2Z	H5M	H5Z
<i>PR</i>	<b>0.8</b>	0.79	<b>0.8</b>	0.79	0.78	0.78	<b>0.8</b>	0.79	<b>0.8</b>	0.79
<i>J</i>	<b>0.21</b>	0.15	0.17	0.13	0.1	0.1	0.19	0.13	0.18	0.12
<i>F</i>	<b>0.37</b>	0.31	0.34	0.29	0.25	0.24	0.36	0.3	0.34	0.28
<i>BGM</i>	<b>0.13</b>	0.1	0.11	0.09	0.07	0.07	<b>0.13</b>	0.1	0.12	0.09
<i>NMI</i>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>	<b>0.13</b>
<i>VI</i>	<b>4.58</b>	5.03	4.77	5.15	5.38	5.47	4.63	5.09	4.69	5.15
<i>LCE</i>	<b>0.22</b>	<b>0.22</b>	0.23	0.23	0.25	0.25	<b>0.22</b>	<b>0.22</b>	<b>0.22</b>	<b>0.22</b>
<i>GCE</i>	0.33	<b>0.3</b>	0.33	0.31	0.34	0.33	0.32	0.31	0.32	0.3
<i>V</i>	0.95	0.95	0.95	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	0.95	0.95	0.95	0.95
<i>MSE</i>	<b>353.64</b>	316.37	359.05	319.85	399.69	384.1	328.79	310.97	330.48	<b>293.36</b>
<i>Z</i>	<b>0.43</b>	0.39	0.32	0.31	0.17	0.17	<b>0.45</b>	0.39	0.39	0.36
<i>B</i>	<b>0.1</b>	0.12	0.11	0.12	0.15	0.16	<b>0.1</b>	0.13	<b>0.1</b>	0.12
Combinaison supremum										
	AM	AZ	GM	GZ	H1M	H1Z	H2M	H2Z	H5M	H5Z
<i>PR</i>	<b>0.8</b>	0.79	<b>0.8</b>	0.79	0.78	0.78	<b>0.8</b>	0.79	<b>0.8</b>	0.79
<i>J</i>	0.19	0.13	0.17	0.12	0.1	0.09	<b>0.2</b>	0.13	0.17	0.12
<i>F</i>	<b>0.36</b>	0.29	0.34	0.28	0.25	0.24	<b>0.37</b>	0.29	0.34	0.28
<i>BGM</i>	<b>0.13</b>	0.09	0.11	0.08	0.07	0.07	<b>0.13</b>	0.09	0.11	0.08
<i>NMI</i>	0.13	0.13	0.13	0.13	0.12	0.13	<b>0.13</b>	0.13	0.13	<b>0.14</b>
<i>VI</i>	4.63	5.03	4.78	5.19	5.35	5.48	<b>4.52</b>	5.12	4.69	5.19
<i>LCE</i>	<b>0.22</b>	<b>0.22</b>	0.23	0.23	0.25	0.25	<b>0.22</b>	<b>0.22</b>	<b>0.22</b>	<b>0.22</b>
<i>GCE</i>	0.32	<b>0.3</b>	0.33	0.31	0.34	0.33	<b>0.33</b>	<b>0.3</b>	0.32	<b>0.3</b>
<i>V</i>	0.95	0.95	0.95	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.95</b>	0.95	0.95	0.95
<i>MSE</i>	349.31	311.45	352.68	314.73	403.86	380.2	<b>331.59</b>	303.01	325.83	<b>287.82</b>
<i>Z</i>	0.43	0.38	0.33	0.31	0.17	0.18	<b>0.45</b>	0.39	0.4	0.36
<i>B</i>	<b>0.1</b>	0.12	0.11	0.12	0.15	0.16	<b>0.1</b>	0.13	<b>0.1</b>	0.12

TABLE 2.5 – Performances moyennes (calculées sur 200 images de la base de BERKELEY avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture) des segmentations par combinaison adaptative et par combinaison supremum [Cohen 2009].

Le tableau 2.5 présente l'évaluation moyenne des segmentations basées sur les combinaisons adaptative et supremum des informations de couleur et de texture et en fonction de chaque métrique d'évaluation, sur 200 images de la base de BERKELEY. À partir de ce tableau, nous constatons que pour la combinaison adaptative, le couple de

gradients AM (gradient basé sur le filtre ASF pour la texture et le gradient morphologique pour la couleur) obtient la première position avec la majorité des métriques d'évaluation. Cependant, pour la combinaison supremum, le couple de gradients H2M (gradient basé sur le 2<sup>ème</sup> paramètre d'HARALICK pour la texture et le gradient morphologique pour la couleur) se trouve être la meilleure combinaison avec la majorité des métriques d'évaluation. Nous constatons aussi que dans le cas de la combinaison adaptative, toutes les combinaisons couleur/texture testées ont été pénalisées de la même manière lorsque l'évaluation est réalisée avec la métrique  $NMI$ .

Le tableau 2.6 présente la performance moyenne (sur 200 images de la base de BERKELEY BSD500 [Arbelaez 2011]) de chaque méthode de segmentation décrite dans le chapitre 1. Les mesures ont été obtenues avec les métriques d'évaluation décrites dans ce chapitre. Les performances moyennes des méthodes de segmentation basées respectivement sur la combinaison fixe, adaptative et supremum ont été réinjectées dans le tableau 2.6, en ne retenant que les valeurs obtenues avec les meilleurs combinaisons couleur/texture (ayant été favorisées par la majorité des métriques d'évaluation utilisées : H2M pour la combinaison fixe, AM pour la combinaison adaptative et H2M pour la combinaison supremum). Notons que les valeurs en rouge (respectivement en bleu) sont les premiers meilleurs résultats (respectivement les deuxièmes meilleurs résultats). Nous pouvons constater que la méthode de segmentation par CTB (Color, Texture and Brightness) basée sur la combinaison des informations de couleur, de texture et de luminance se positionne comme la meilleure méthode de segmentation pour la plupart des métriques d'évaluation (4 fois en première position et 3 fois en deuxième position). La deuxième meilleure méthode de segmentation est la segmentation basée sur la combinaison fixe (CF) (3 fois en première position et 1 fois en deuxième position). La troisième meilleure méthode de segmentation est la segmentation par MeanShift (MS) (2 fois en première position et 2 fois en deuxième position). Nous constatons aussi qu'une segmentation effectuée avec l'algorithme JSEG se positionne toujours en dernier, hormis avec la métrique  $MSE$  où le JSEG se place en première position et avec le critère de VINET où il se place en deuxième position. Nous pouvons conclure également que le critère de VINET peut être considéré comme étant moins approprié en comparaison avec les autres métriques utilisées puisqu'il n'arrive pas à départager sans équivoque la performance des différentes méthodes de segmentation (une segmentation en première position, sept en deuxième et le reste en troisième position).

Critère d'évaluation	Méthodes de segmentation										
	MS	JSEG	FH	FCR	NCUTS	CTM	CV	CTB	CF	CA	CS
Évaluation supervisée											
$PR$	0.73	0.75	0.81	0.76	0.78	0.73	0.76	0.79	0.8	0.8	0.8
$J$	0.34	0.09	0.36	0.32	0.25	0.33	0.29	0.38	0.29	0.21	0.2
$F$	0.51	0.19	0.52	0.49	0.41	0.5	0.45	0.55	0.45	0.37	0.37
$BGM$	0.3	0.07	0.24	0.25	0.14	0.28	0.24	0.28	0.24	0.13	0.13
$NMI$	0.06	0.04	0.11	0.08	0.1	0.06	0.08	0.09	0.13	0.13	0.13
$VI$	4.15	6.06	4.0	3.86	4.05	4.05	5.21	1.47	3.96	4.58	4.52
$LCE$	0.29	0.56	0.26	0.28	0.22	0.26	0.31	0.23	0.18	0.22	0.22
$GCE$	0.32	0.63	0.37	0.34	0.32	0.31	0.41	0.27	0.29	0.33	0.33
$V$	0.95	0.95	0.94	0.95	0.95	0.94	0.94	0.95	0.96	0.95	0.95
Évaluation non-supervisée											
$MSE$	285.26	252.42	380.25	288.86	506.38	478.57	429.82	392.72	328.18	353.64	331.59
$Z$	0.61	0.12	0.52	0.28	0.33	0.29	0.28	0.35	0.48	0.43	0.45
$B$	0.07	0.09	0.13	0.05	0.18	0.07	0.2	0.06	0.10	0.10	0.10

TABLE 2.6 – Performances moyennes des principales méthodes de segmentation de la littérature sur 200 images de la base de BERKELEY : MS MeanShift, JSEG, FH Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR, CTM Compression-based Texture Merging, CV segmentation de CHAN-VESE, CTB segmentation basée sur la combinaison de la couleur, la texture et la luminosité (Color, Texture and Brightness), CF segmentation basée sur la combinaison fixe, CA segmentation basée sur la combinaison adaptative et CS segmentation basée sur la combinaison supremum.

À présent, nous faisons appel à la base de texture VISTEX et nous mesurons la performance de la segmentation basée sur une combinaison fixe des informations de couleur et de texture, en fonction des valeurs du paramètre de pondération  $\alpha$  et des différents couples de gradients couleur et texture utilisés. Le tableau 2.7 présente pour chaque combinaison fixe couleur/texture et en fonction du paramètre  $\alpha$ , la performance de la segmentation mesurée avec les critères d'évaluation non-supervisée (sans segmentation de référence)  $MSE$ ,  $\mathcal{Z}$  de ZEBOUDJ et  $\mathcal{B}$  de BORSOTTI.

AM				AZ				GM			
$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$
0	658.73	0.14	0.2	0	658.73	0.14	0.2	0	645.09	0.03	0.19
0.2	530.12	0.29	0.18	0.2	478.84	0.31	0.19	0.2	516.5	0.1	0.18
0.5	413.51	0.4	<b>0.16</b>	0.5	374.95	0.37	<b>0.18</b>	0.5	428.12	0.29	0.17
0.7	<b>376.57</b>	0.42	<b>0.16</b>	0.7	<b>374.75</b>	0.38	<b>0.18</b>	0.7	<b>378.38</b>	0.38	<b>0.15</b>
1	385.49	<b>0.45</b>	<b>0.16</b>	1	436.41	<b>0.39</b>	0.2	1	385.49	<b>0.45</b>	0.16

GZ				H1M				H1Z			
$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$
0	645.09	0.03	0.19	0	541.17	0.02	0.24	0	541.17	0.02	0.24
0.2	487.57	0.11	0.19	0.2	561.24	0.04	0.25	0.2	565.03	0.05	0.26
0.5	393.55	0.29	<b>0.18</b>	0.5	442.63	0.18	0.21	0.5	428.48	0.2	0.2
0.7	<b>358.43</b>	0.36	<b>0.18</b>	0.7	<b>371.16</b>	0.28	0.18	0.7	<b>365.25</b>	0.27	<b>0.19</b>
1	436.41	<b>0.39</b>	0.2	1	385.49	<b>0.45</b>	<b>0.16</b>	1	436.41	<b>0.39</b>	0.2

H2M				H2Z				H5M			
$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$	$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$
0	633.73	0.24	0.17	0	633.73	0.24	<b>0.17</b>	0	489.95	0.19	0.16
0.2	527.55	0.32	0.16	0.2	473.19	0.35	<b>0.17</b>	0.2	457.11	0.26	<b>0.14</b>
0.5	429.69	0.43	<b>0.15</b>	0.5	388.12	<b>0.4</b>	<b>0.17</b>	0.2	410.74	0.38	<b>0.14</b>
0.7	398.64	<b>0.45</b>	<b>0.15</b>	0.7	<b>372.59</b>	<b>0.4</b>	0.18	0.7	<b>384.62</b>	0.42	0.15
1	<b>385.49</b>	<b>0.45</b>	0.16	1	436.41	0.39	0.2	1	385.49	<b>0.45</b>	0.16

H5Z			
$\alpha$	$MSE$	$\mathcal{Z}$	$\mathcal{B}$
0	489.95	0.19	<b>0.16</b>
0.2	446.27	0.28	<b>0.16</b>
0.5	<b>363.3</b>	0.37	<b>0.16</b>
0.7	375.08	0.38	0.18
1	436.41	<b>0.39</b>	0.2

TABLE 2.7 – Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec les métriques d'évaluation non-supervisée  $MSE$ ,  $\mathcal{Z}$  de ZEBOUDJ et  $\mathcal{B}$  de BORSOTTI) de la méthode de segmentation basée sur une combinaison fixe couleur/texture en fonction des gradients couleur/texture et du paramètre de pondération  $\alpha$ .

Dans le tableau 2.8, nous retenons les meilleures valeurs obtenues pour chaque métrique d'évaluation. Nous constatons que le  $MSE$  favorise la valeur  $\alpha^* = 0.7$  pour la plupart des combinaisons couleur/texture hormis avec les combinaisons H2M ( $\alpha^* = 1$ ) et H5Z



( $\alpha^* = 0.5$ ). La mesure  $\mathcal{Z}$  favorise  $\alpha^* = 1$  pour la majorité des combinaisons couleur/texture hormis avec la combinaison H2Z ( $\alpha^* = 0.5, 0.7$ ). La mesure  $\mathcal{B}$  fournit quant à elle plutôt des valeurs hétérogènes. Nous pouvons donc conclure que les deux mesures  $\mathcal{MSE}$  et  $\mathcal{Z}$  pénalisent les combinaisons donnant plus de poids aux gradients couleur (ie. quand  $\alpha$  est faible).

	Combinaison fixe					
	$\mathcal{MSE}$	$\alpha^*$	$\mathcal{Z}$	$\alpha^*$	$\mathcal{B}$	$\alpha^*$
AM	376.57	0.7	<b>0.45</b>	1	0.16	0.5, 0.7, 1
AZ	374.75	0.7	0.39	1	0.18	0.5, 0.7
GM	378.38	0.7	<b>0.45</b>	1	0.15	0.7
GZ	<b>358.43</b>	0.7	0.39	1	0.18	0.5, 0.7
H1M	371.16	0.7	<b>0.45</b>	1	0.16	1
H1Z	365.25	0.7	0.39	1	0.19	0.7
H2M	385.49	1	<b>0.45</b>	0.7, 1	0.15	0.5, 0.7
H2Z	372.59	0.7	0.4	0.5, 0.7	0.17	0, 0.2, 0.5
H5M	<b>384.62</b>	0.7	<b>0.45</b>	1	<b>0.14</b>	0.2, 0.5
H5Z	363.3	0.5	0.39	1	0.16	0, 0.2, 0.5

TABLE 2.8 – Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec  $\mathcal{MSE}$ ,  $\mathcal{Z}$  et  $\mathcal{B}$ ) de la méthode de segmentation basée sur la combinaison fixe couleur/texture en fonction de la valeur optimale  $\alpha^*$  du paramètre de combinaison.

	Combinaison adaptative				Combinaison supremum		
	$\mathcal{MSE}$	$\mathcal{Z}$	$\mathcal{B}$		$\mathcal{MSE}$	$\mathcal{Z}$	$\mathcal{B}$
AM	429.66	0.4	0.17	AM	410.4	0.4	0.16
AZ	376.1	0.36	0.18	AZ	373.92	0.36	0.18
GM	442.63	0.31	0.17	GM	438.88	0.3	0.17
GZ	404.05	0.31	0.18	GZ	397.86	0.31	0.18
H1M	432.42	0.17	0.21	H1M	441.34	0.18	0.21
H1Z	407.74	0.18	0.21	H1Z	408.15	0.19	0.21
H2M	<b>388.96</b>	<b>0.41</b>	<b>0.16</b>	H2M	<b>392.85</b>	<b>0.41</b>	<b>0.15</b>
H2Z	372.84	0.38	0.18	H2Z	362.79	0.39	0.18
H5M	422.28	0.37	<b>0.16</b>	H5M	416.66	0.38	0.16
H5Z	<b>362.89</b>	0.36	0.18	H5Z	<b>357.9</b>	0.36	0.18

TABLE 2.9 – Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec  $\mathcal{MSE}$ ,  $\mathcal{Z}$  et  $\mathcal{B}$ ) des méthodes de segmentation basées sur les combinaisons adaptative et supremum des informations de couleur et de texture.

Tout comme nous l'avons fait pour la segmentation par combinaison fixe couleur/texture, nous mesurons la performance des segmentations basées sur les combinaisons adaptative et supremum en fonction du couple gradient couleur/texture sachant que le paramètre de pondération est déterminé de manière adaptative en fonction du contenu de l'image. L'évaluation est présentée dans le tableau 2.9. Les mesures ont été réalisées sur 33 images de la base de texture VISTEX, choisies aléatoirement. Nous constatons que la mesure  $\mathcal{MSE}$  favorise le couple H5Z (5<sup>ème</sup> paramètre d'HARALICK pour la texture et le gradient de DI-ZENZO pour la couleur) pour les combinaisons adaptative et supremum, tandis que les deux autres mesures  $\mathcal{Z}$  et  $\mathcal{B}$  favorisent le couple H2M (2<sup>ème</sup> paramètre



d'HARALICK pour la texture et morphologique pour la couleur).

Méthode de segmentation	$MSE$	$\mathcal{Z}$	$\mathcal{B}$
MS	218.65	0.55	0.06
Jseg	449.82	0.15	0.08
FH	422.61	0.52	0.16
FCR	350.8	0.33	0.08
Ncuts	667.2	0.27	0.21
CTM	671.79	0.22	0.07
CV	532.33	0.29	0.24
CF	358.43	0.45	0.14
CA	362.89	0.41	0.16
CS	357.9	0.41	0.15

TABLE 2.10 – Comparatif (sur 33 images de texture de la base de VisTex) des performances moyennes des principales méthodes de segmentation de la littérature .

Le tableau 2.10 et la figure 2.21 présentent une étude comparative des différentes méthodes de segmentation de la littérature (décrites dans le chapitre 1). Les performances moyennes des méthodes de segmentation basées respectivement sur la combinaison fixe, adaptative et supremum ont été réinjectées dans le tableau 2.10 en ne retenant que les valeurs obtenues avec les meilleurs combinaisons couleur/texteure qui ont été favorisées par la majorité des métriques d'évaluation (H5M pour la combinaison fixe et H2M pour les combinaisons adaptative et supremum). Le tableau 2.10 résume donc les performances moyennes (sur 33 images de texture sélectionnées aléatoirement à partir de la base de texture VisTex) de chaque méthode de segmentation. Notons que, dans le tableau 2.10, les valeurs en rouge (respectivement en bleu) sont les meilleurs résultats (respectivement les deuxièmes meilleurs résultats). La figure 2.21 illustre également les performances de ces méthodes de segmentation que ce soit individuellement (pour chaque image) ou en moyenne. Les méthodes de segmentation ont été évaluées avec les métriques d'évaluation non-supervisée ( $MSE$ ,  $\mathcal{Z}$  et  $\mathcal{B}$ ) puisque les images de la base de VisTex ne possèdent pas de segmentations de référence. D'après le tableau 2.10 et la figure 2.21, nous concluons que le MEANSHIFT offre les segmentations les plus performantes sur la base de texture VISTEX. En effet, il obtient les meilleures valeurs quelle que soit la métrique d'évaluation utilisée. La méthode de segmentation qui se place en deuxième position après le MEANSHIFT est la FCR. La CTM se positionne directement après la méthode de segmentation FH.

Les figures 2.22 et 2.23 illustrent visuellement les résultats de segmentation obtenus par les méthodes testées et évaluées dans ce chapitre sur une sélection de huit images couleur extraites de la base de BERKELEY. Pour une meilleure visualisation, nous n'avons illustré pour la combinaison couleur/texteure fixe [Angulo 2007], que les résultats obtenus avec le couple gradients couleur/texteure optimal H2M mais pour différentes valeurs de  $\alpha$  (paramètre de pondération). Les résultats de segmentation par combinaisons adaptative et supremum ont été également illustrés pour les couples gradients couleur/texteure optimaux (qui sont respectivement le couple AM et le couple H2M). Sur le même principe, les figures 2.24 et 2.25 illustrent les résultats de segmentation sur une sélection de cinq images de texture extraites de la base VisTex.

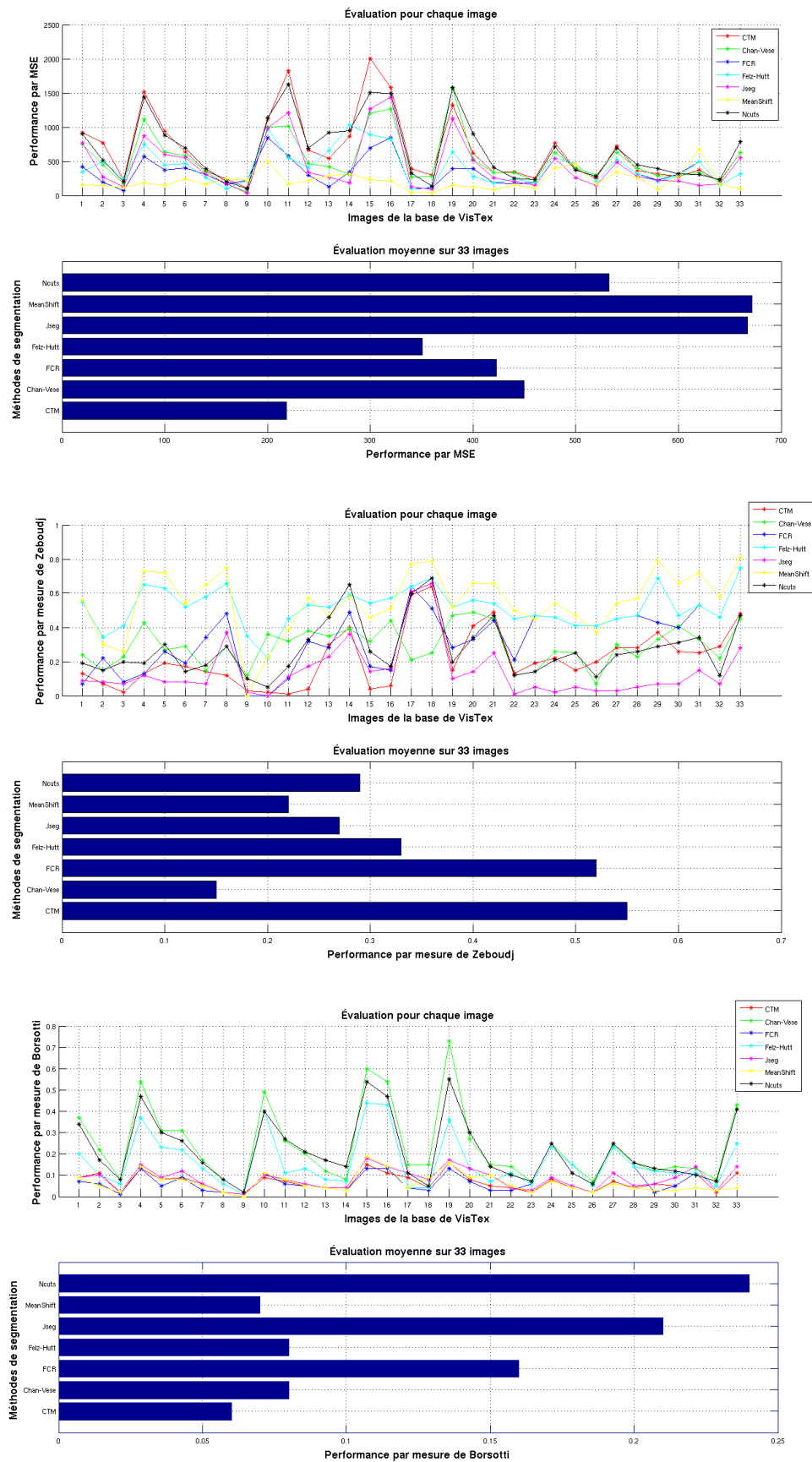


FIGURE 2.21 – Évaluation (sur 33 images de la base de VisTex) des méthodes de segmentation couleur/texture pour chaque image et en moyenne.

Au regard des segmentations obtenues avec les différentes méthodes, nous pouvons confirmer que ces résultats qualitatifs coïncident avec les résultats quantitatifs donnés auparavant. En effet, il apparaît que la méthode de segmentation CTB basée sur la combinaison des informations couleur, texture et luminance (2<sup>ème</sup> ligne de la figure 2.22) produit les meilleurs résultats avec une meilleure préservation des contours des objets présents dans chaque image.

## 2.5/ CONCLUSION

Face à la diversité des méthodes de segmentation, comme présenté dans le chapitre précédent, il est important d'évaluer chaque méthode afin de déterminer celle qui se trouve la plus adaptée aux besoins de l'utilisateur. C'est la raison pour laquelle, nous avons effectué dans ce chapitre, une étude comparative des méthodes de segmentation les plus fréquemment utilisées dans la littérature. L'évaluation a été réalisée à l'aide de métriques d'évaluation supervisée ( $PR$ ,  $J$ ,  $F$ ,  $F$ ,  $NMI$ ,  $VI$ ,  $LCE$ ,  $GCE$  et  $V$ ) et non-supervisées ( $MSE$ ,  $Z$  et  $B$ ). Dans le cas d'une évaluation supervisée, chaque image analysée doit disposer au minimum d'une segmentation de référence fournie par un expert. Deux bases d'images souvent utilisées dans la littérature pour le test des méthodes de segmentation ont été utilisées : une base d'images couleur générique connue sous le nom de base de BERKELEY [Arbelaez 2011] (où chacune des images possède au moins cinq segmentations de référence) et une base de texture connue sous le nom de VISTEX [VisTex 1995].

À partir de l'étude comparative effectuée dans ce chapitre, nous avons conclu que deux méthodes de segmentation ont donné les meilleures performances. La première méthode est la segmentation basée sur la CTB (Colour, Texture and Brightness) proposée dans les travaux de ARBELAEZ et al. [Arbelaez 2011]. En effet, elle a produit les meilleures valeurs de la majorité des métriques utilisées. La deuxième méthode est la segmentation par MEANSHIFT [Comaniciu 2002]. Les méthodes de segmentation basées sur la combinaison des informations de couleur et de texture (fixe [Angulo 2007], adaptative et supremum [Cohen 2009]) ont aussi réalisé des résultats satisfaisants en considérant le couple de gradients couleur/texture optimal.

L'étude comparative réalisée dans ce chapitre servira à examiner la performance de la méthode de segmentation que nous proposons dans le chapitre suivant. Ceci nous permettra de pouvoir la comparer puis la positionner par rapport aux principales méthodes existantes dans la littérature.

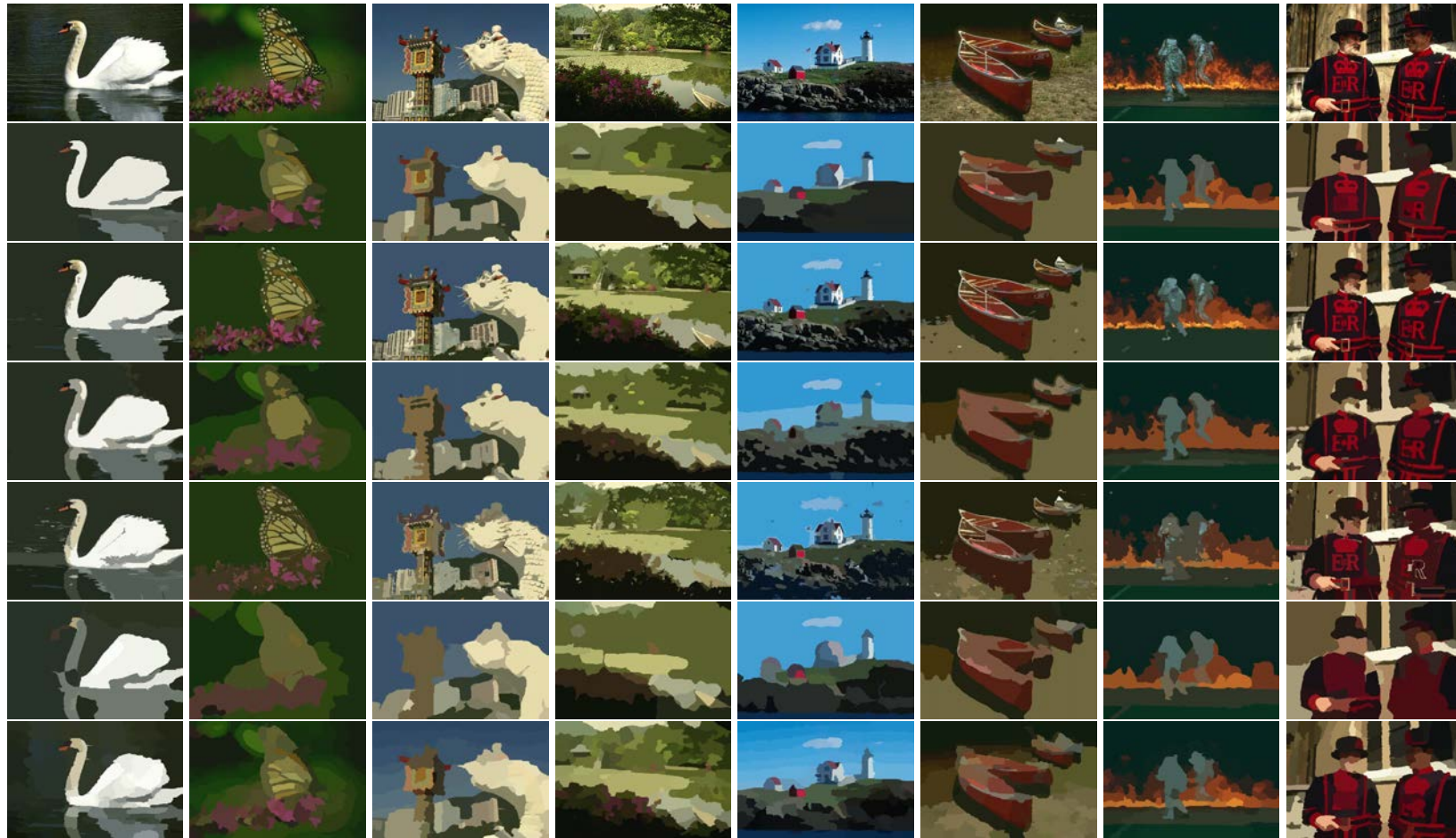


FIGURE 2.22 – Illustration des segmentations sur huit images de la base de BERKELEY [Arbelaez 2011] (de haut en bas : les images originales, images segmentées par CTB, images segmentées par MeanShift, images segmentées par FCR, images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, images segmentées par CTM et images segmentées par Jseg).



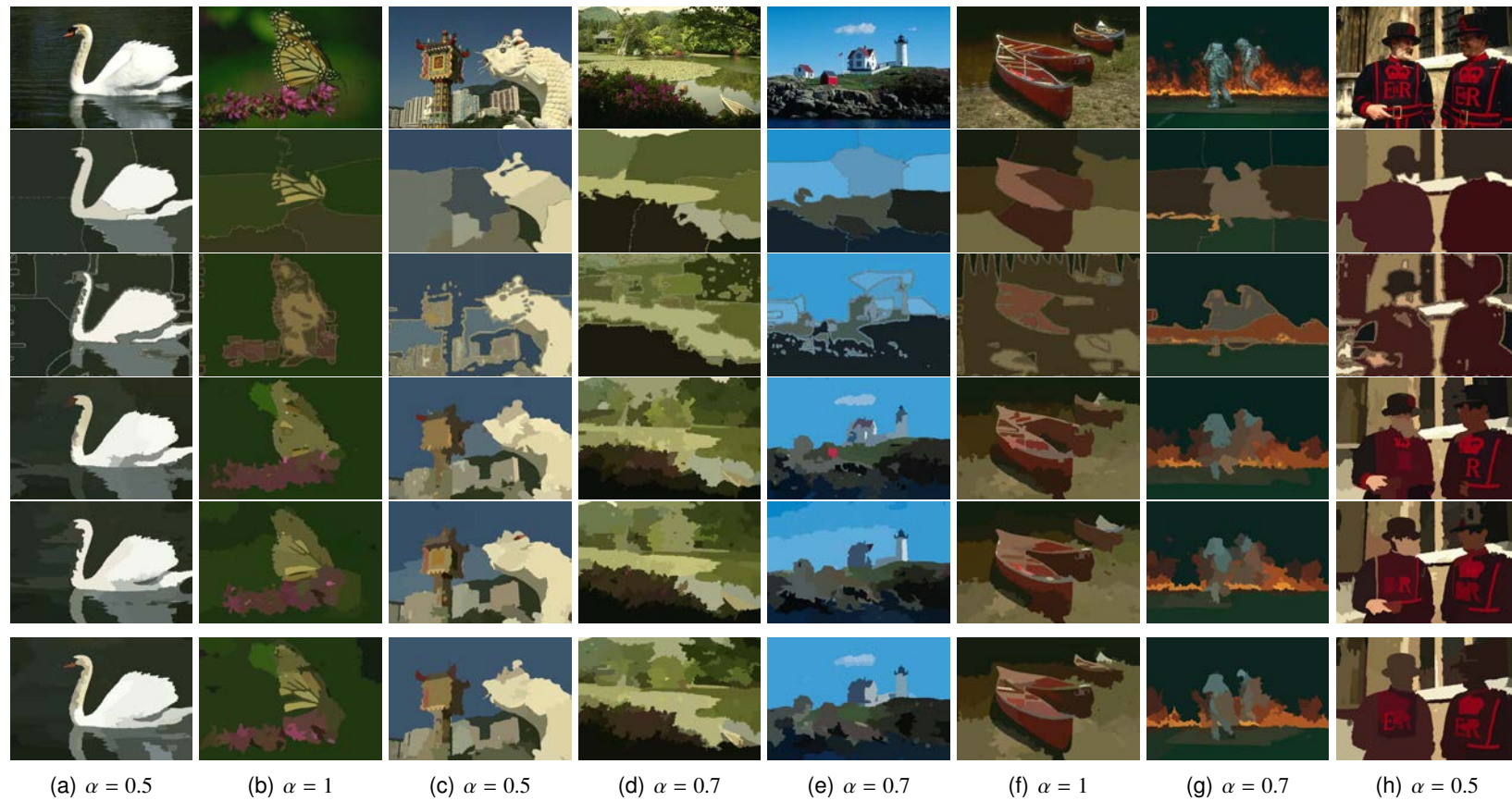


FIGURE 2.23 – Illustration des segmentations sur huit images de la base de BERKELEY [Arbelaez 2011] (de haut en bas : les images originales, images segmentées par Ncuts, images segmentées par CHAN-VESE, images segmentées par combinaison adaptative, images segmentées par combinaison supremum et images segmentées par combinaison fixe).

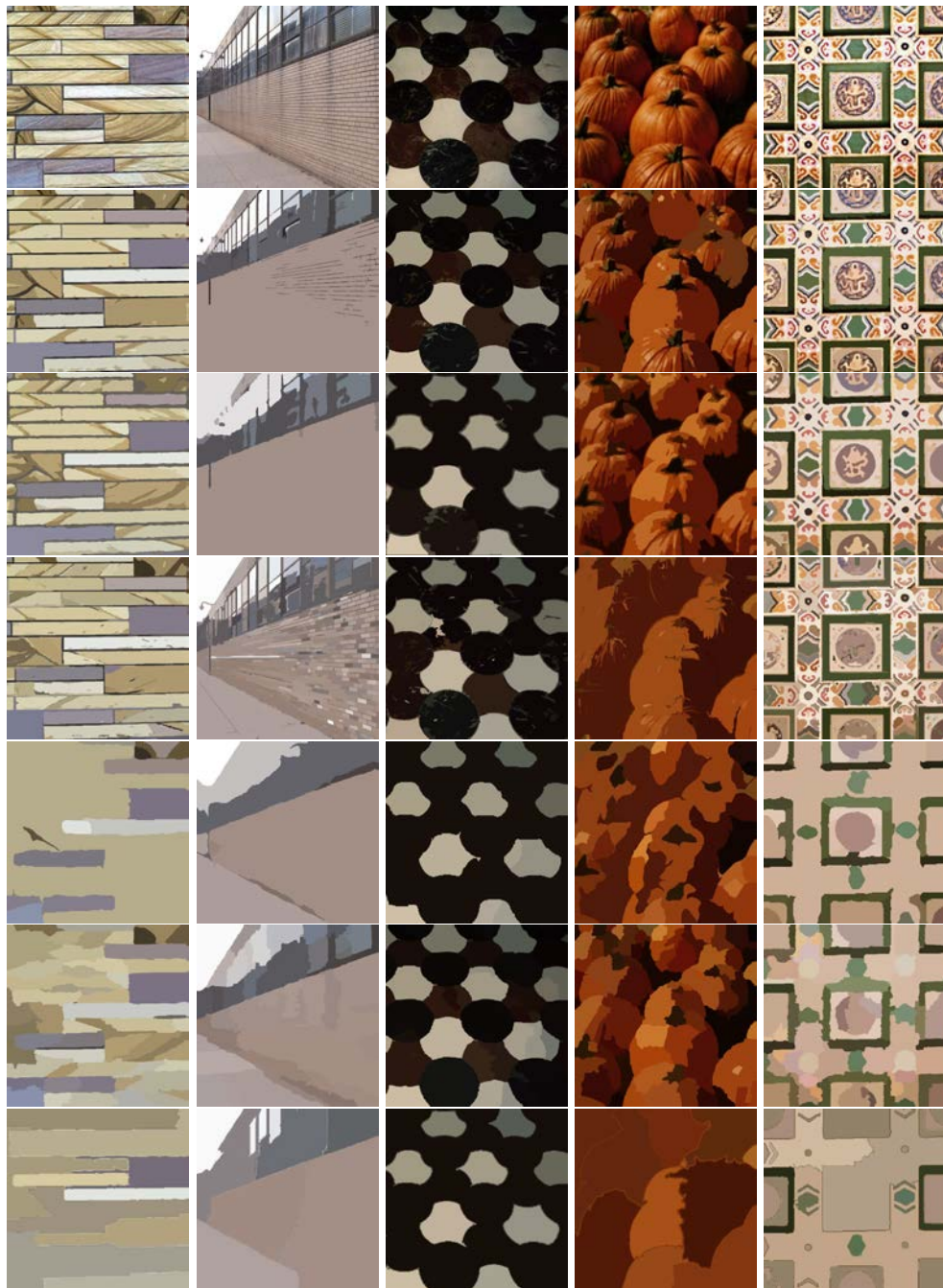


FIGURE 2.24 – Illustration des segmentations sur cinq images de la base de texture VisTex [VisTex 1995](de haut en bas : les images originales, images segmentées par MeanShift, images segmentées par FCR, images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, images segmentées par CTM, images segmentées par Jseg et images segmentées par Ncuts).



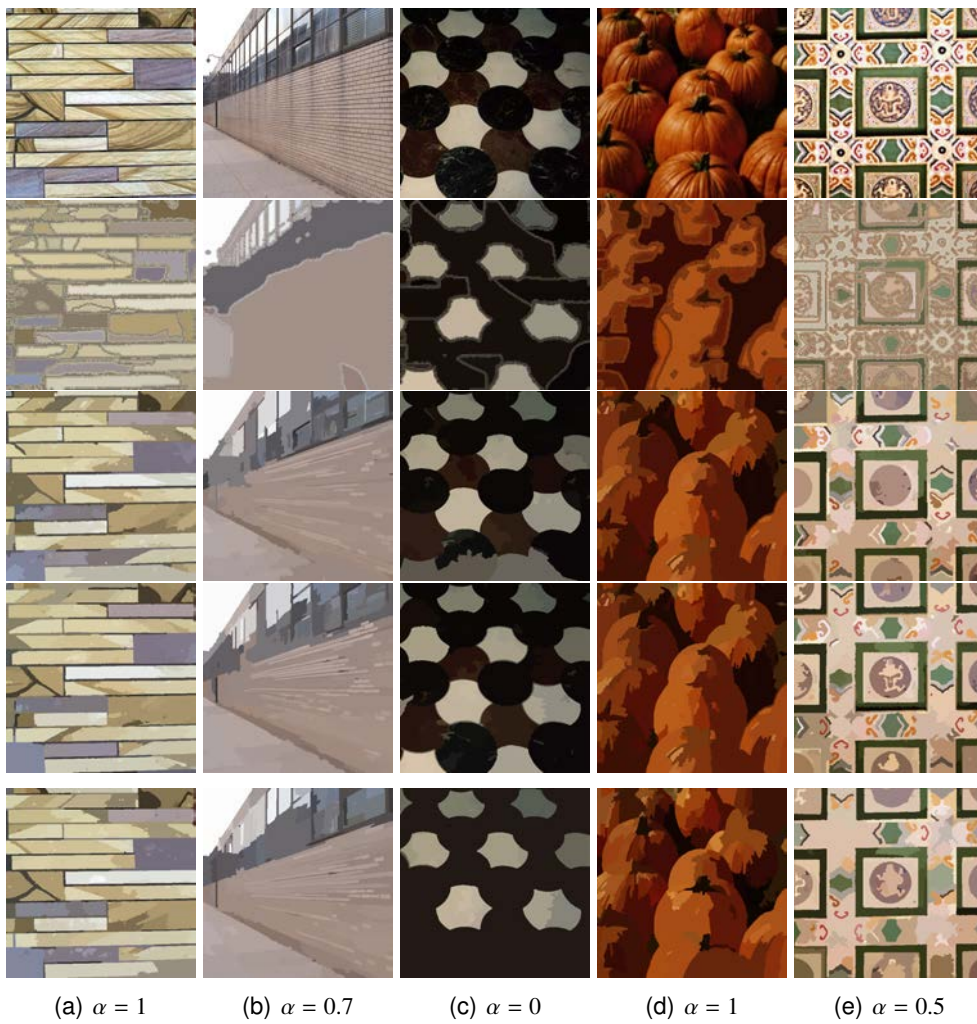


FIGURE 2.25 – Illustration des segmentations sur cinq images de la base de texture VisTex [VisTex 1995] (de haut en bas : les images originales, images segmentées par CHAN-VESE, images segmentées par combinaison adaptative, images segmentées par combinaison supremum et images segmentées par combinaison fixe).



# UNE MÉTHODE DE COMBINAISON COULEUR-TEXTURE NON PARAMÉTRIQUE

## Sommaire

<b>3.1</b>	<b>Introduction</b>	<b>101</b>
<b>3.2</b>	<b>Les descripteurs de couleur et de texture</b>	<b>102</b>
3.2.1	Les descripteurs de couleur	103
3.2.2	Les descripteurs de texture	105
<b>3.3</b>	<b>Une nouvelle méthode non paramétrique de combinaison couleur/texture</b>	<b>117</b>
3.3.1	L'analyse en composantes principales	117
3.3.2	La combinaison couleur/texture par comparaison de valeurs propres	120
<b>3.4</b>	<b>La méthode de Ligne de Partage des Eaux</b>	<b>124</b>
<b>3.5</b>	<b>Résultats expérimentaux</b>	<b>127</b>
<b>3.6</b>	<b>Conclusion</b>	<b>140</b>

## 3.1/ INTRODUCTION

Dans le premier chapitre, nous avons présenté un état de l'art exhaustif des méthodes de segmentation d'images utilisant conjointement les informations de couleur et de texture. Nous avons introduit trois méthodes de segmentation basées sur une combinaison paramétrique de ces deux sources d'information : la combinaison fixe proposée dans les travaux d'ANGULO [Angulo 2007] et les combinaisons adaptative et supremum proposées dans les travaux de COHEN et al. [Cohen 2009]. L'objectif de ce chapitre est de proposer une nouvelle méthode adaptative et non paramétrique pour la combinaison des informations couleur/texture. Cette méthode est basée sur l'analyse spectrale et l'étude des axes principaux d'un ensemble de points 3D formé à partir de descripteurs couleur et texture calculés sur l'image à analyser. L'intérêt de la combinaison couleur/texture proposée est de ne retenir que les informations couleur et texture les plus pertinentes, à partir d'un nombre donné (choisi par l'utilisateur) de descripteurs.

Le présent chapitre est organisé comme suit : nous rappelons tout d'abord les définitions des principaux descripteurs couleur et texture que nous avons utilisés pour effectuer nos expérimentations, sachant que la méthode de combinaison couleur/texture proposée peut

utiliser un ensemble de descripteurs plus large. Nous détaillons ensuite, le fondement et les hypothèses de l'approche de combinaison proposée, en rappelant le principe de l'Analyse en Composantes Principales (ACP). Dans un troisième temps, nous présentons une méthode de segmentation issue de la morphologie mathématique, bien connue de la littérature pour ses très bons résultats, à savoir la Ligne de Partage des Eaux (LPE). Le gradient hybride obtenu à l'aide de la combinaison couleur/texture proposée constituera une entrée de l'algorithme de LPE que nous utiliserons pour segmenter l'image. En dernière partie, nous effectuons une étude qualitative et quantitative des résultats de segmentation basée sur la nouvelle combinaison couleur/texture, en utilisant les métriques d'évaluation introduites dans le deuxième chapitre. Nous situons également les performances de cette méthode par rapport aux autres méthodes de segmentation de la littérature présentées dans le chapitre 1.

### 3.2/ LES DESCRIPTEURS DE COULEUR ET DE TEXTURE

La nouvelle méthode de combinaison couleur/texture que nous proposons et développons dans ce chapitre consiste à partir d'un ensemble de descripteurs de ne retenir que les informations couleur et texture les plus pertinentes dans l'image. Dans cette thèse, nous avons défini comme descripteurs, des gradients couleur et texture générés à partir de différents attributs couleur et texture décrits dans le premier chapitre. En ce qui concerne les gradients couleur, nous avons choisi le gradient de DI-ZENZO et le gradient morphologique couleur (avec un ordre lexicographique). En ce qui concerne les gradients texture, une classification des attributs de texture issus de la littérature a été effectuée dans le premier chapitre. Cette classification distingue cinq familles d'attributs que nous rappelons ci-après. Peu utilisée, la première famille regroupe les attributs géométriques. La deuxième famille concerne les attributs basés sur la modélisation spatiale tels que les modèles auto-régressifs, les modèles de Markov et les motifs locaux binaires (Local Binary Patterns : LBP). La troisième famille rassemble les attributs basés sur la modélisation spatio-fréquentielle telles que la transformée de Gabor, la transformée en ondelettes et la transformée de Fourier. La quatrième famille concerne les attributs statistiques tels que les attributs de premier ordre (la moyenne et la variance), les attributs de second ordre (les histogrammes et les matrices de co-occurrences) et les attributs d'ordre supérieur (les matrices de longueurs de plages). Finalement, la cinquième famille regroupe les attributs issus de la morphologie mathématique.

Dans le but d'obtenir la meilleure description du contenu de l'image (couleur et texture), nous avons eu recours aux attributs les plus fréquemment utilisés dans la littérature. Robustes et efficaces, ces attributs ne se contentent pas seulement de réaliser une modélisation spatiale mais permettent aussi de donner une représentation fréquentielle. C'est la raison pour laquelle nous nous intéressons au filtre de Gabor. Les matrices de co-occurrences constituent également un outil souvent utilisé, ayant prouvé son efficacité dans la description de l'information texture présente dans l'image. Nous retenons également le filtre alterné séquentiel (Alternate Sequential Filter : ASF) qui est un attribut issu de la morphologie mathématique. Nous utilisons ce filtre composite et complexe car il permet de décrire la texture de façon beaucoup plus fine que les autres attributs (ouverture et fermeture par exemple) de la morphologie mathématique.

À présent, nous détaillons ci-dessous les gradients calculés sur les attributs couleur/texture. Ces gradients interviendront dans les tests de la combinaison couleur/texture proposée. Dans le cadre de cette thèse, les expérimentations ont été réalisées avec un nombre précis de descripteurs. Notons que ce nombre peut bien évidemment être étendu grâce au mode de calcul de la combinaison proposée.

### 3.2.1/ LES DESCRIPTEURS DE COULEUR

Dans cette section, nous allons définir les descripteurs couleur que nous avons sélectionnés, à savoir le gradient de DI-ZENZO et le gradient morphologique couleur.

#### 3.2.1.1/ GRADIENT MORPHOLOGIQUE COULEUR

Le gradient morphologique est un opérateur issu de la théorie de la morphologie mathématique [Najman 2008]. Ce gradient a été défini par BEUCHER dans le cas discret [Beucher 1990]. D'autres définitions du gradient morphologique (directionnel, régularisé, thik/thin, etc.) ont été présentées dans les travaux de RIVEST et al. [Rivest 1993]. L'expression de calcul la plus courante du gradient morphologique est celle proposée par BEUCHER. Elle est donnée dans le cas discret par l'équation 3.1. Il s'agit du résidu du dilaté d'une image  $I$  et de son érodé.

$$\mathcal{G}_M(I) = \delta_{ES}(I) - \varepsilon_{ES}(I) \quad (3.1)$$

où  $ES$  est un élément structurant donné.  $\delta_{ES}$  et  $\varepsilon_{ES}$  sont respectivement les opérations de dilatation et d'érosion.

Dans notre cas, le gradient morphologique a été calculé sur l'image couleur initiale en utilisant un élément structurant  $ES$  carré (en 8-connexité), de taille  $3 \times 3$  pixels et un ordre lexicographique. La figure 3.1 présente un exemple de gradient morphologique couleur calculé sur une image couleur de la base de BERKELEY.

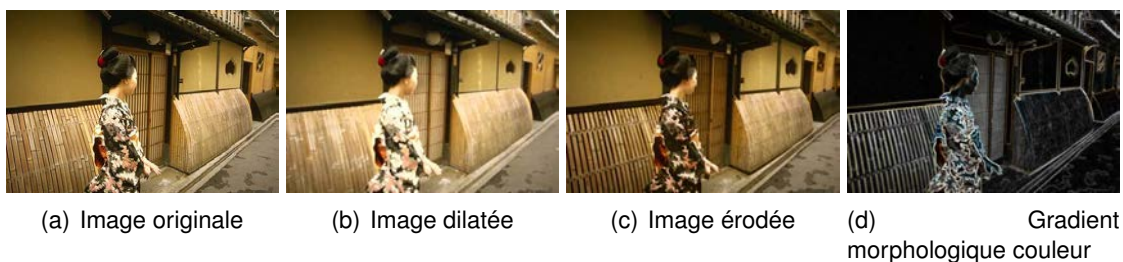


FIGURE 3.1 – Illustration du gradient morphologique couleur calculé sur une image de la base de BERKELEY.

#### 3.2.1.2/ GRADIENT DE DI-ZENZO

Le gradient de DI-ZENZO est un gradient multi-composantes s'appliquant sur des images couleur [DiZenko 1986, Lambert 2002]. Son mode de calcul est composé de deux

étapes. La première consiste à calculer séparément sur chaque composante couleur, la dérivée première de l'image. Après cette étape, nous disposons de deux images : une image d'amplitude et une image de direction du gradient (pour chaque composante couleur de l'image). L'amplitude du gradient est obtenue par le maximum de la dérivée en un pixel de l'image de coordonnées  $x$  et  $y$  tandis que la direction est obtenue par l'équation  $\text{artang}(\frac{dy}{dx})$ . La valeur d'amplitude du gradient en un point reflète la variation de niveau de gris observée dans l'image en ce point. Cette variation est considérée forte quand la valeur de l'amplitude est élevée. La direction du gradient est orthogonale à la frontière qui passe par le point considéré. La deuxième étape du calcul du gradient de DI-ZENZO consiste à fusionner ces informations selon un critère choisi a priori.

Dans la pratique, l'algorithme de DI-ZENZO est basé sur la recherche des valeurs propres de la matrice  $M$  donnée par les équations suivantes [Dubois 2008] :

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

tels que

$$a = \mathcal{G}_{x_1}^2 + \mathcal{G}_{x_2}^2 + \mathcal{G}_{x_3}^2 \quad (3.2)$$

$$b = \mathcal{G}_{x_1} \times \mathcal{G}_{y_1} + \mathcal{G}_{x_2} \times \mathcal{G}_{y_2} + \mathcal{G}_{x_3} \times \mathcal{G}_{y_3} \quad (3.3)$$

$$c = \mathcal{G}_{y_1}^2 + \mathcal{G}_{y_2}^2 + \mathcal{G}_{y_3}^2 \quad (3.4)$$

avec  $\mathcal{G}_{x_1}$ ,  $\mathcal{G}_{x_2}$  et  $\mathcal{G}_{x_3}$  (respectivement  $\mathcal{G}_{y_1}$ ,  $\mathcal{G}_{y_2}$  et  $\mathcal{G}_{y_3}$ ) représentent le gradient selon le déplacement  $dx$  (respectivement  $dy$ ) sur chacune des trois composantes. Les valeurs propres de la matrice  $M$  sont utilisées pour calculer le module du gradient (cf. équation 3.5) et les vecteurs propres associés traduisent les directions du gradient  $\theta_{\mathcal{G}}$  (cf. équation 3.6).

$$\|\mathcal{G}\| = \sqrt{\lambda_1^{\mathcal{G}} + \lambda_2^{\mathcal{G}}} \quad (3.5)$$

avec

$$\begin{aligned} \lambda_1^{\mathcal{G}} &= \frac{1}{2} \times (a + c + \sqrt{(a - c)^2 - 4 \times b^2}) \\ \lambda_2^{\mathcal{G}} &= \frac{1}{2} \times (a + c - \sqrt{(a - c)^2 - 4 \times b^2}) \\ \theta_{\mathcal{G}} &= \frac{1}{2} \times \arctan\left(\frac{2 \times b}{(a - c)}\right) \end{aligned} \quad (3.6)$$

Dans notre cas, le gradient de DI-ZENZO a été calculé sur l'image couleur initiale en utilisant des codes de FREEMAN dans l'intervalle  $[0, \dots, 7]$ . La figure 3.2 présente un exemple de calcul du gradient de DI-ZENZO sur une image couleur issue de la base de BERKELEY.

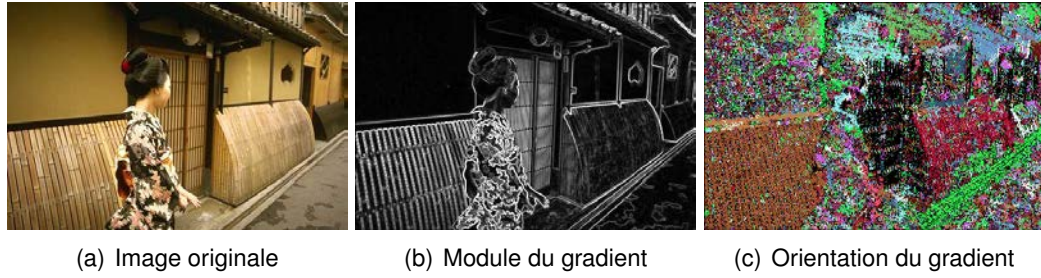


FIGURE 3.2 – Illustration du gradient de DI-ZENZO sur une image couleur de la base de BERKELEY.

### 3.2.2/ LES DESCRIPTEURS DE TEXTURE

Les gradients de texture que nous décrivons dans cette section sont basés sur les attributs de texture faisant partie des attributs texture les plus communément utilisés dans les applications de la segmentation d'images couleur/texture. Ces gradients sont basés sur les filtres de Gabor, des paramètres d'HARALICK extraits des matrices de co-occurrences et le filtre alterné séquentiel.

#### 3.2.2.1/ GRADIENT BASÉ SUR LES FILTRES DE GABOR

La transformation de Gabor est souvent définie comme une transformée de Fourier avec fenêtrage [Palm 2002]. Par conséquent, cette transformation permet d'examiner la régularité de la texture dans le domaine fréquentiel tout en tenant compte de sa localisation spatiale. La transformée de Gabor est obtenue par l'équation 3.8, en appliquant la transformée de Fourier décrite dans les travaux de [Palm 2002] avec un fenêtrage gaussien.

Soient  $I$  une image initiale de taille  $L_x \times L_y$  pixels,  $w(x_\alpha - x_0, y_\alpha - y_0)$  la fenêtre gaussienne entre la position spatiale  $(x_0, y_0)$  et la position spatiale  $(x_\alpha, y_\alpha)$  après une rotation spatiale (d'un angle  $\alpha$ ) du vecteur  $(x, y)^t$  (cf. équation 3.7).

$$\begin{pmatrix} x_\alpha \\ y_\alpha \end{pmatrix} = A_\alpha \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{avec} \quad A_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (3.7)$$

$$\mathcal{T}_{G_{u_0, v_0}}(x_0, y_0) = \frac{1}{L_x L_y} \sum_{x=0}^{L_x-1} \sum_{y=0}^{L_y-1} I(x, y) \cdot w(x_\alpha - x_0, y_\alpha - y_0) \cdot \exp\left(-2\pi i \left(\frac{x u_0}{L_x} + \frac{y v_0}{L_y}\right)\right) \quad (3.8)$$

où  $x$  et  $y$  sont les coordonnées spatiales telles que  $x \in \{0, \dots, L_x - 1\}$  et  $y \in \{0, \dots, L_y - 1\}$ .  $u_0$  et  $v_0$  sont les fréquences spatiales correspondant aux coordonnées spatiales  $x_0$  et  $y_0$ .

Dans le cas général,  $u$  et  $v$  sont les fréquences spatiales des coordonnées spatiales  $x$  et  $y$  telles que  $u \in \{-\frac{L_x}{2}, \dots, \frac{L_x}{2}\}$  et  $v \in \{-\frac{L_y}{2}, \dots, \frac{L_y}{2}\}$ . Elles représentent respectivement la décomposition verticale et horizontale.

En utilisant l'aspect symétrique de la fenêtre de Gauss  $w(x, y)$  (cf. équation 3.9), la transformée de Gabor peut être décrite par l'équation 3.10. Le masque de convolution  $\mathcal{F}_{G_{u_0, v_0, \alpha}(x_0, y_0)}$  représente le filtre de Gabor.

$$w(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \cdot \exp\left(-\frac{1}{2} \frac{x^2}{\sigma_x^2}\right) \cdot \exp\left(-\frac{1}{2} \frac{y^2}{\sigma_y^2}\right) \quad (3.9)$$

$$\mathcal{T}_{G_{u_0, v_0}}(x_0, y_0) = I(x_0, y_0) \times \mathcal{F}_{G_{u_0, v_0, \alpha}(x_0, y_0)} \quad (3.10)$$

La caractérisation de texture d'une image couleur par les filtres de Gabor est obtenue en appliquant ces filtres (sensible chacun à une fréquence donnée) sur chacune des fenêtres balayant les composantes couleur [Palm 2000, Palm 2002]. Dans nos expérimentations,  $\alpha$  prend les valeurs  $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3}{4}\pi\}$  et la taille de la fenêtre de Gauss prend les valeurs  $21 \times 21$  pixels et  $35 \times 35$  pixels. Pour un filtre donné, l'attribut de texture peut être obtenu en calculant l'énergie, l'entropie ou la variance des coefficients de Gabor générés pour chaque fenêtre. Les expressions de ces quantités sont données dans les travaux de PALM et al. [Palm 2002]. Dans cette thèse, nous avons choisi l'énergie comme attribut de texture étant le paramètre le plus utilisé dans la littérature. Le descripteur de texture basé sur les filtres de Gabor est alors obtenu en calculant le gradient morphologique sur chaque couche de texture. À l'issue de cette étape, nous retenons, pour chaque pixel, la valeur maximale parmi toutes les couches de texture. La figure 3.3 présente une illustration des résultats (sur les trois composantes couleur) du gradient basé sur les filtres de Gabor.

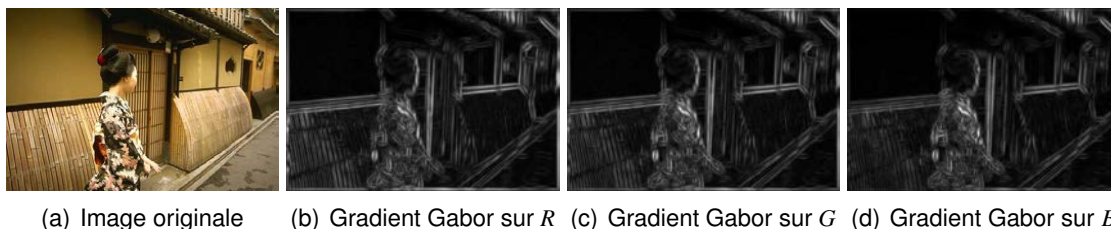


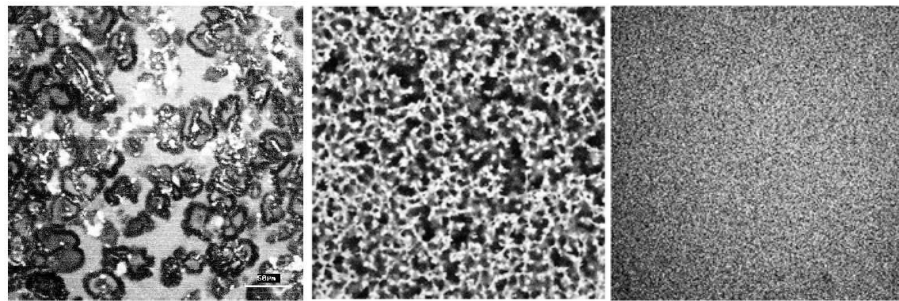
FIGURE 3.3 – Illustration du descripteur de texture basé sur les filtres de Gabor sur une image couleur issue de la base de BERKELEY dans l'espace couleur  $(R, G, B)$ .

### 3.2.2.2/ GRADIENTS BASÉS SUR LES PARAMÈTRES D'HARALICK

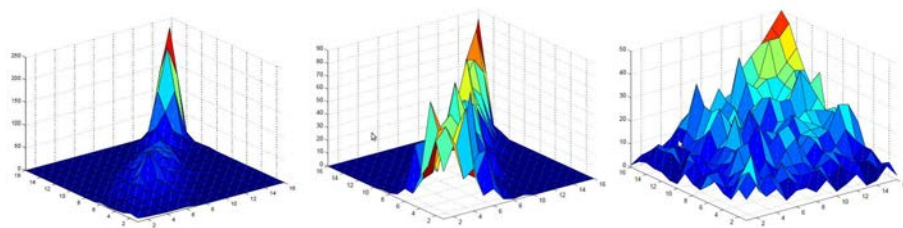
Les matrices de co-occurrences proposées par HARALICK en 1973 [Haralick 1973], constituent un outil performant pour la description des interactions spatiales entre les pixels. En effet, cet outil permet d'étudier la distribution de l'information couleur dans le cas des matrices de co-occurrences chromatiques tout en tenant compte de la relation spatiale entre les pixels de l'image à analyser. Les matrices de co-occurrences d'une image  $I$  sont calculées à l'aide de l'équation 1.2 (déjà introduite dans le chapitre 1). Dans la majorité des cas, les matrices de co-occurrences ne sont pas directement exploitées pour le calcul des attributs de texture. Ceci s'explique par la difficulté d'interpréter visuellement des surfaces de co-occurrences. À titre d'illustration, la figure 3.4 présente trois échantillons d'images de textures de gel [Hibon 2010]. La seule information que nous pouvons déduire de cette illustration est que pour les textures grossières, les matrices de co-occurrences ont tendance à avoir des valeurs



élevées près de la diagonale principale, tandis que pour les textures fines, les valeurs sont dispersées. Ceci montre l'apport insuffisant de l'exploitation directe des surfaces de co-occurrences. Par ailleurs, les matrices de co-occurrences mettent en évidence d'énorme quantité d'informations. C'est pourquoi les utilisateurs extraient de ces matrices des attributs moins coûteux en mémoire. Ainsi, 14 attributs connus sous le nom de paramètres d'HARALICK [Haralick 1973] ont été introduits afin de conserver la pertinence des matrices de co-occurrences.



(a) Images initiales de textures de gel.



(b) Matrices de co-occurrences pour  $\|\vec{d}\| = 1$  et  $\theta = 0$ .

FIGURE 3.4 – Illustration des matrices de co-occurrences de textures de gel calculées sur une fenêtre de  $60 \times 60$  pixels sur 16 niveaux de gris ( $N = 16$ ) [Hibon 2010].

Dans le cas pratique, les 14 paramètres d'HARALICK ne sont pas tous utilisés. Seulement 6 paramètres (l'énergie, le contraste, la corrélation, la variance, l'homogénéité et l'entropie), davantage interprétables, sont souvent exploités pour le calcul d'attributs de texture. Il reste difficile de donner, même de manière intuitive, une signification précise pour les autres paramètres [Haralick 1973, Porebski 2009]. Nous présentons ci-après les 14 paramètres d'HARALICK en illustrant sur une image de la base de BERKELEY, ceux qui restent les plus performants et donc les plus utilisés dans la littérature.

Soient  $I$  l'image initiale,  $m$  et  $n$  deux niveaux de gris de  $I$ ,  $C_k$  et  $C'_k$  deux composantes couleur de  $I$  et  $\vec{d}$  le vecteur de déplacement.

#### a) Second moment angulaire (Énergie/Uniformité)

Le second moment angulaire (Second Angular Moment : SAM), appelé aussi énergie, s'agit d'une mesure qui examine l'homogénéité globale de l'image. Cet attribut atteint de fortes valeurs lorsque la distribution des niveaux de gris est constante ou de forme périodique. Dans le cas particulier où les deux niveaux de gris  $m$  et  $n$  sont égaux, il s'agit d'une mesure d'uniformité. Dans ce cas, le paramètre SAM est d'autant plus faible que les valeurs non-nulles de la diagonale de la matrice de co-occurrences sont équi-réparties. Plus l'indice est élevé, plus le nombre de niveaux de gris renvoyant à



eux-mêmes par la translation  $\vec{d}$  est faible [Sharma 2001a]. L'expression du paramètre SAM s'écrit par l'équation donnée ci-après :

$$H_1^{C_k, C_{k'}, \vec{d}} = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left( M^{C_k C_{k'}} [I](\vec{d})(m, n) \right)^2 \quad (3.11)$$

#### b) Contraste

Le contraste mesure les variations locales des niveaux de gris (couleurs) présents dans l'image à analyser. Cette mesure est importante lors du passage (réciproque) d'un niveau faible à un niveau élevé. Une valeur faible de cet attribut caractérise une texture plutôt uniforme [Sharma 2001b]. Le contraste s'écrit par l'équation donnée ci-après :

$$H_2^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{N-1} l^2 \left( \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m, n) \right) \quad (3.12)$$

#### c) Entropie

La mesure d'entropie comme son nom l'indique examine le désordre que peut présenter l'image [Sharma 2001a]. Plus précisément, l'entropie caractérise la dispersion spatiale et spectrale des valeurs non-nulles de la matrice de co-occurrences. Cette mesure prend des valeurs faibles lorsque les couples de pixels sont bien répartis dans l'image et des valeurs importantes lorsque ces couples sont en désordre. L'entropie s'exprime par l'équation donnée ci-après :

$$H_3^{C_k, C_{k'}, \vec{d}} = - \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m, n) \log (M^{C_k C_{k'}} [I](\vec{d})(m, n)) \quad (3.13)$$

#### d) Corrélation

La corrélation permet de mesurer les dépendances en ligne et en colonne entre les niveaux de composantes couleur dans l'image [Haralick 1973, Paquis 2003]. L'expression de la corrélation est décrite par l'équation donnée ci-après :

$$H_4^{C_k, C_{k'}, \vec{d}} = \frac{(m - \mu_m^M)(n - \mu_n^M) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m, n)}{\sigma_m^M \sigma_n^M} \quad (3.14)$$

où  $\mu_m^M$ ,  $\mu_n^M$ ,  $\sigma_m^M$  et  $\sigma_n^M$  constituent respectivement les centres de gravité et les écarts type de  $M_m^{C_k C_{k'}} [I](\vec{d})(m)$  et de  $M_n^{C_k C_{k'}} [I](\vec{d})(n)$ , avec :

$$M_m^{C_k C_{k'}} [I](\vec{d})(m) = \sum_{n=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m, n) \quad (3.15)$$

et

$$M_n^{C_k C_{k'}} [I](\vec{d})(n) = \sum_{m=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m, n) \quad (3.16)$$

### e) Variance (Inertie)

Le cinquième paramètre d'HARALICK est la variance, appelé aussi inertie. Intuitivement, cette mesure permet d'examiner la distribution de chaque composante couleur autour de sa moyenne [Majdoulayne 2009]. Autrement dit, la variance mesure l'hétérogénéité de la texture et augmente lorsque les niveaux de gris diffèrent de leur moyenne. Elle est décrite par l'équation donnée ci-après :

$$H_5^{C_k, C_{k'}, \vec{d}} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} (m - \mu^M)^2 M^{C_k C_{k'}} [I](\vec{d})(m, n) \quad (3.17)$$

où  $\mu^M$  est le centre de gravité des coefficients de la matrice  $M^{C_k C_{k'}} [I]$ .

### f) Moment des différences inverses (homogénéité)

Le moment différentiel inverse correspond à une mesure d'homogénéité locale au sein de la texture [Paquis 2003]. En effet, l'homogénéité traduit la forte ou la faible occurrence d'un couple de pixels donné, séparé par l'isométrie  $\vec{d}$ . Cette mesure est élevée dans les zones de l'image où le niveaux de gris est uniforme ou lorsque l'image présente des motifs de périodes  $\|\vec{d}\|$  dans la direction  $\theta$ . L'expression de ce paramètre est décrite par l'équation donnée ci-après :

$$H_6^{C_k, C_{k'}, \vec{d}} = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \frac{1}{1 + (m - n)^2} M^{C_k C_{k'}} [I](\vec{d})(m, n) \quad (3.18)$$

Afin d'être exhaustif, nous citons l'expression des autres paramètres d'HARALICK, peu utilisés du fait de leur interprétation difficile.

### g) Moyenne des sommes

$$H_7^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{2(N-1)} l \cdot M_{m+n}^{C_k C_{k'}} [I](\vec{d})(l) \quad (3.19)$$

où

$$M_{m+n}^{C_k C_{k'}} [I](\vec{d})(l) = \sum_{\substack{m'=0 \\ m'+n'=l}}^{N-1} \sum_{n'=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m', n')$$

### h) Variance des sommes

$$H_8^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{2(N-1)} (l - H_7^{C_k, C_{k'}, \vec{d}}) \cdot M_{m+n}^{C_k C_{k'}} [I](\vec{d})(l) \quad (3.20)$$

### i) Entropie des sommes

$$H_9^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{2(N-1)} M_{m+n}^{C_k C_{k'}} [I](\vec{d})(l) \log (M_{m+n}^{C_k C_{k'}} [I](\vec{d})(l)) \quad (3.21)$$

j) **Variance des différences**

$$H_{10}^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{N-1} (1 - \mu_{m-n}^M) M_{m-n}^{C_k C_{k'}} [I](\vec{d})(l) \quad (3.22)$$

où

$$M_{m-n}^{C_k C_{k'}} [I](\vec{d})(l) = \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m', n') \quad l = 0, 1, \dots, N-1$$

et

$$\mu_{m-n}^M = \sum_{l=0}^{N-1} l. M_{m-n}^{C_k C_{k'}} [I](\vec{d})(l)$$

k) **Entropie des différences**

$$H_{11}^{C_k, C_{k'}, \vec{d}} = \sum_{l=0}^{N-1} M_{m-n}^{C_k C_{k'}} [I](\vec{d})(l) \log (M_{m-n}^{C_k C_{k'}} [I](\vec{d})(l)) \quad (3.23)$$

l) **Informations sur la corrélation**

$$H_{12}^{C_k, C_{k'}, \vec{d}} = \frac{H_3^{C_k, C_{k'}, \vec{d}} - H_{mn}^1}{\max(H_m^1, H_n^1)} \quad (3.24)$$

où

$$H_{mn}^1 = - \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} M^{C_k C_{k'}} [I](\vec{d})(m', n') \log (M_m^{C_k C_{k'}} [I](\vec{d})(m') \times M_n^{C_k C_{k'}} [I](\vec{d})(n'))$$

$$H_m^1 = - \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} M_m^{C_k C_{k'}} [I](\vec{d})(m') \log (M_m^{C_k C_{k'}} [I](\vec{d})(m'))$$

$$H_n^1 = - \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} M_n^{C_k C_{k'}} [I](\vec{d})(n') \log (M_n^{C_k C_{k'}} [I](\vec{d})(n'))$$

m) **Informations sur la corrélation**

$$H_{13}^{C_k, C_{k'}, \vec{d}} = \left( 1 - \exp[-2.(H_{mn}^2 - H_3^{C_k, C_{k'}, \vec{d}})] \right)^{\frac{1}{2}} \quad (3.25)$$

où

$$H_{mn}^2 = - \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} M_m^{C_k C_{k'}} [I](\vec{d})(m') \times M_n^{C_k C_{k'}} [I](\vec{d})(n') \log (M_m^{C_k C_{k'}} [I](\vec{d})(m') \times M_n^{C_k C_{k'}} [I](\vec{d})(n'))$$

#### n) Coefficient de corrélation maximale

$$H_{14}^{C_k, C_{k'}, \vec{d}} = (\text{Plus grande valeur propre de } Q)^{\frac{1}{2}} \quad (3.26)$$

où

$$Q(\vec{d})(m, n) = \sum_{l=0}^{N-1} \frac{M^{C_k C_{k'}}[I](\vec{d})(m, l) \times M^{C_k C_{k'}}[I](\vec{d})(n, l)}{M_m^{C_k C_{k'}}[I](\vec{d})(m) \times M_n^{C_k C_{k'}}[I](\vec{d})(n)}$$

Dans cette thèse, les gradients basés sur les paramètres d'HARALICK ont été calculés en appliquant un gradient morphologique sur les images d'attributs de texture  $H_h$  (énergie, contraste et variance), comme décrit dans l'équation 3.27. Les matrices de co-occurrences ont été générées en balayant l'image avec une fenêtre (de taille  $4 \times 4$  pixels) et avec un niveau de quantification  $N = 16$ . Plusieurs combinaisons de translations ont été testées ( $\|\vec{d}\| \in \{3, 5\}$  pixels) dans différentes directions  $\theta \in \{0, \dots, 150\}$ .

$$\mathcal{G}_{H_h} = \delta_{ES}(H_h) - \varepsilon_{ES}(H_h) \quad , \forall h \in \{1, \dots, 14\} \quad (3.27)$$

où  $\delta_{ES}$  et  $\varepsilon_{ES}$  sont respectivement les opérations de dilatation et d'érosion en utilisant l'élément structurant  $ES$ .

La figure 3.5 illustre trois gradients calculés à partir des trois paramètres d'HARALICK que nous jugeons opportuns (énergie, contraste et variance), sur une image de la base de BERKELEY. En effet, pour une meilleure visualisation nous n'avons illustré que les gradients dont le résultat est exploitable.

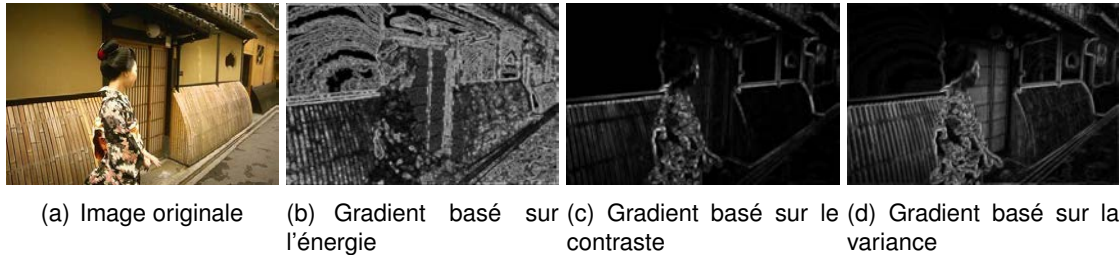


FIGURE 3.5 – Gradients calculés à partir d'attributs d'HARALICK.

Au vu de la pertinence et l'apport des attributs de texture calculés à partir des matrices de co-occurrences, d'autres paramètres ont été introduits, à savoir le moment diagonal, la différence inverse, la dissimilarité, le troisième moment angulaire, le cluster shade, etc. Ces paramètres sont décrits dans les travaux [Anys 1995, Porebski 2009, Bombardier 2009]. Dans cette thèse, nous nous contentons des paramètres d'HARALICK.

#### 3.2.2.3/ GRADIENT BASÉ SUR LE FILTRE ALTERNÉ SÉQUENTIEL

Le filtre alterné séquentiel (Alternate Sequential Filter : ASF) est un outil issu de la morphologie mathématique. Il consiste à appliquer une succession de fermetures  $\varphi$  et d'ouvertures  $\gamma$  avec des éléments structurants  $\{ES_q\}_{q \in 1, \dots, o}$  de taille croissante jusqu'à

un ordre  $o$  donné [Soille 1999]. Ce filtre permet de simplifier l'image en conservant ses détails (sans lissage), ce qui permet de caractériser la texture en préservant les contrastes. Posons  $I$  l'image initiale, le filtre alterné séquentiel d'ordre  $p$ , que nous notons  $\mathcal{ASF}_o(I)$ , est défini par l'équation suivante :

$$\mathcal{ASF}_o(I) = \varphi_{ES_o}(\gamma_{ES_o}(\dots \varphi_{ES_2}(\gamma_{ES_2}(\varphi_{ES_1}(\gamma_{ES_1}(I)))))) \quad (3.28)$$

La figure 3.6 illustre la simplification d'une image issue de la base de BERKELEY par un filtre ASF avec différents ordres.

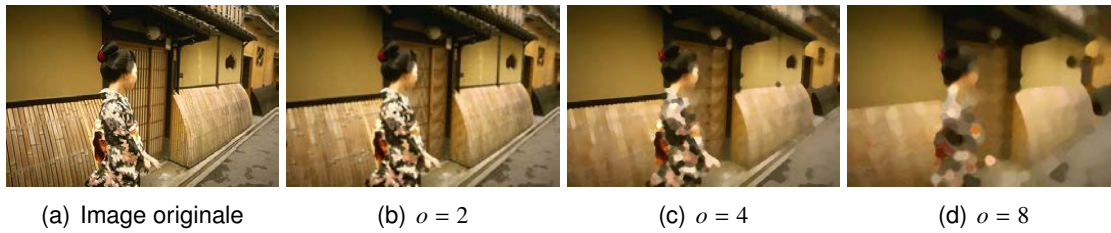


FIGURE 3.6 – Illustration d'une simplification d'une image issue de la base de BERKELEY par un filtre ASF avec différents ordres.

Pour le calcul d'un gradient de texture basé sur un filtre ASF, nous avons suivi l'approche proposée par ANGULO [Angulo 2007] qui consiste à redéfinir ce qu'est une couche de texture en utilisant une analyse multi-échelle locale, puis à introduire une nouvelle définition d'un gradient de texture basé sur la morphologie mathématique. Ce gradient morphologique de texture a été calculé à l'aide d'opérations issues de la morphologie mathématique dont nous détaillons brièvement le principe composé de quatre étapes : le filtrage de l'image initiale, la génération d'une couche de texture, l'analyse granulométrique et finalement le calcul du gradient de texture morphologique.

#### a) Filtrage de l'image initiale

La première étape de calcul du gradient morphologique de texture consiste à appliquer un filtrage sur l'image initiale afin de supprimer les détails les moins significatifs tout en conservant la forme et les contours des objets. Le filtre choisi est un filtre connexe de type nivellement. Ce filtre, issu de la morphologie mathématique, permet de conserver les zones homogènes et de supprimer les pics du bruit dans l'image. Soient  $I$  l'image initiale,  $I_N$  l'image filtrée par le filtre de nivellement et  $I_{Marqueur}$  une image de marqueur donnée. La valeur en un pixel  $p$  de l'image filtrée peut être obtenue par itération jusqu'à idempotence (stabilité du résultat) des opérations données par l'équation 3.29.

$$N(p) = \begin{cases} \min(I(p), \delta_{ES}(I_{Marqueur})(p)) & \text{si } I_{Marqueur}(p) < I(p) \\ \max(I(p), \varepsilon_{ES}(I_{Marqueur})(p)) & \text{si } I_{Marqueur}(p) > I(p) \end{cases} \quad (3.29)$$

où  $\delta_{ES}$  est la dilatation morphologique et  $\varepsilon_{ES}$  est l'érosion morphologique avec un élément structurant  $ES$ . L'image  $I_{Marqueur}$  est souvent le résultat de la simplification de l'image initiale  $I$ . Dans notre cas,  $I_{Marqueur}$  est la simplification de  $I$  par un filtre  $\mathcal{ASF}$ .

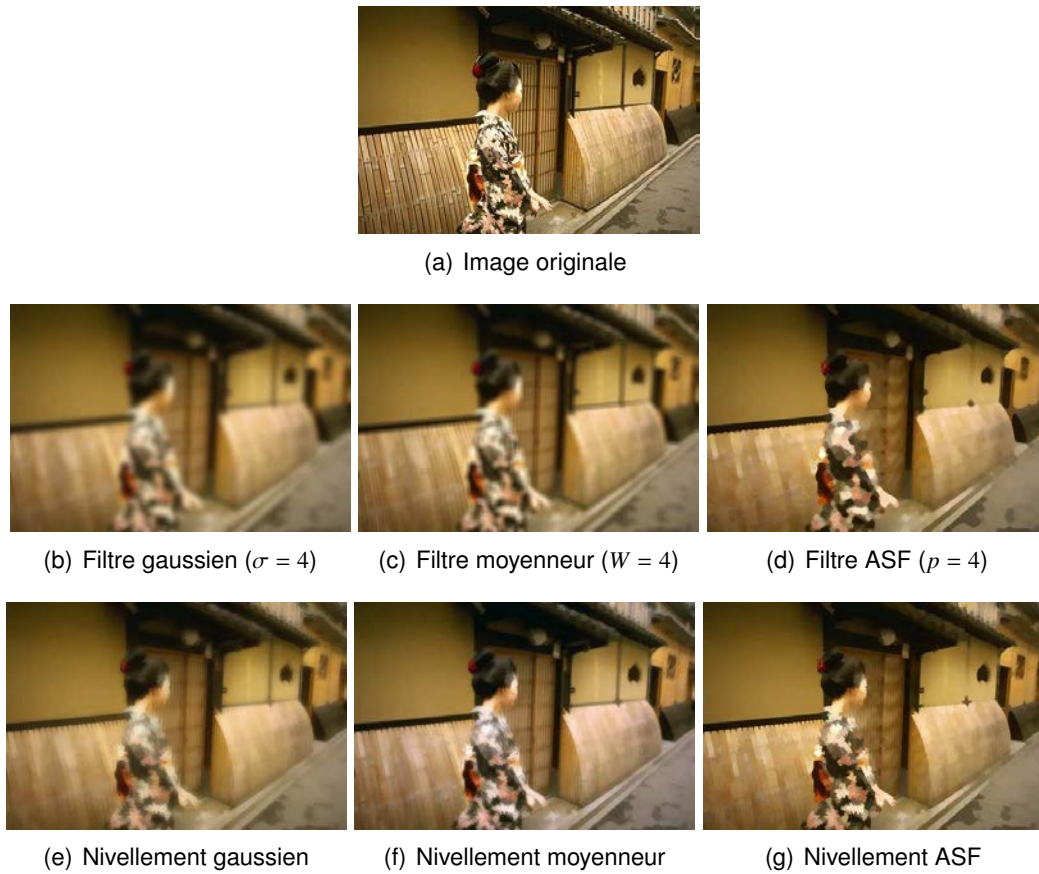


FIGURE 3.7 – Illustration d'un filtrage de nivellement sur une image de la base de BERKELEY bruitée avec un bruit gaussien (de moyenne 5 et d'écart type 10) et avec trois images marqueurs (obtenues avec un filtre gaussien, un filtre moyeneur et un filtre ASF).

La figure 3.7 illustre les différents résultats obtenus en fonction de l'image de marqueur utilisée. Nous pouvons constater que le filtre de nivellement couplé au filtrage par ASF offre la meilleure simplification de l'image tout en conservant les contours des objets.

## b) Définition de la couche de texture

ANGULO [Angulo 2007], tout comme SOFOU et al. [Sofou 2005], ont défini la texture par le résidu de nivellement de la composante de luminance. En effet, ANGULO affirme que le phénomène de texture dans l'image n'est qu'une variation d'intensité lumineuse. Soit  $I^L$  la composante de luminance de l'image initiale  $I^{HSL}$  présentée dans l'espace couleur  $(H, S, L)$  et  $I_N^L$  la composante de luminance de l'image résultante du filtrage par nivellement. La formulation mathématique de l'obtention de la couche de texture telle que définie par ANGULO est donnée par l'équation 3.30 et illustrée dans la figure 3.8.

$$I_{tex} = I^L - I_N^L \quad (3.30)$$



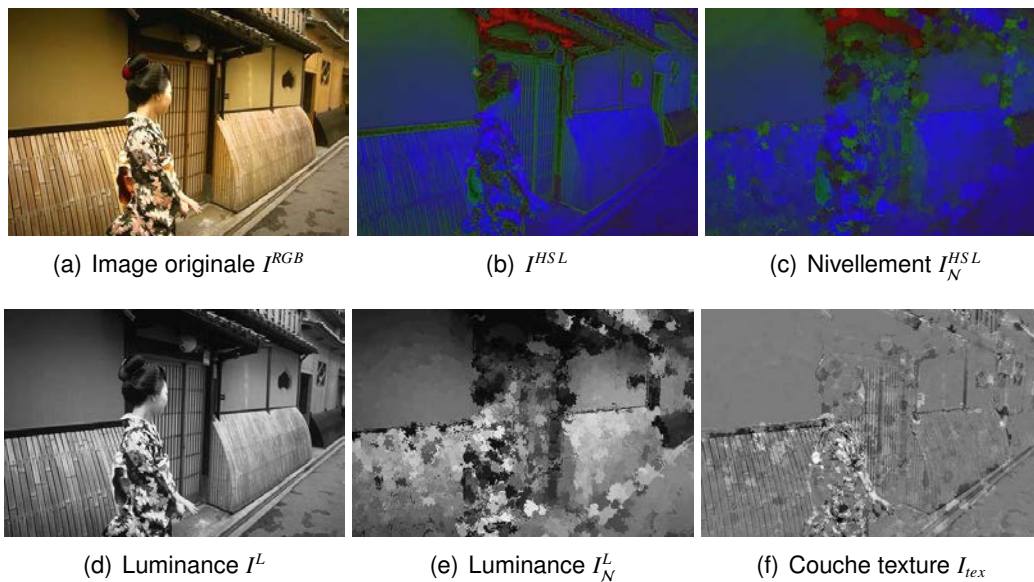


FIGURE 3.8 – Calcul de la couche de texture.

### c) Analyse granulométrique

La granulométrie morphologique est une technique permettant l'étude des tailles des objets dans l'image en se basant sur le principe de tamisage [Angelini 2007]. Cette technique permet la sélection successive d'objets de tailles données croissantes, ce qui donne une idée sur la distribution de leur taille dans l'image. Formellement, l'analyse granulométrique consiste à appliquer un ensemble de tamis de tailles croissantes. Ces tamis sont des transformations  $(\Psi_\eta)_{\eta \geq 0}$  vérifiant les propriétés ci-dessous et ne dépendant que d'un seul paramètre  $\eta \geq 0$ . Ces transformations sont souvent des ouvertures morphologiques  $(\gamma_\eta)_{\eta \geq 0}$  de tailles croissantes.

- $\Psi_0(I) = I$
- $\forall \eta \geq 0, \forall I_1, I_2 \subset I, \Psi_\eta$  est croissante :  $I_1 \subset I_2 \Rightarrow \Psi_\eta(I_1) \subset \Psi_\eta(I_2)$
- $\forall \eta \geq 0, \forall I_1 \subset I, \Psi_\eta$  est anti-extensive :  $\Psi_\eta(I_1) \subset I_1$
- Loi d'absorption :  $\forall \eta \geq 0, \forall \varkappa \geq 0$ , étant deux tailles différentes,  $\Psi_\eta \Psi_\varkappa = \Psi_{\max(\eta, \varkappa)} = \Psi_{\varkappa} \Psi_\eta$

En considérant les transformations granulométriques comme étant une série d'ouvertures  $(\gamma_\eta)_{\eta \geq 0}$  de tailles croissantes, l'analyse granulométrique consiste donc à évaluer chaque ouverture de taille  $\eta$  avec une mesure  $S$  définie comme la somme des valeurs des pixels de l'ouverture (cf. équation 3.31). Une série de fermetures de tailles croissantes  $(\varphi_\eta)_{\eta \geq 0}$  constitue donc une anti-granulométrie (granulométrie pour les objets sombres dans l'image  $I$ )

$$S(\gamma_\eta(I)) = \sum_{p \in \gamma_\eta(I)} \gamma_\eta(I)(p) \quad (3.31)$$

L'ouverture de taille  $\eta$ , étant composée que par les structures claires de  $I$  de taille supérieure à  $\eta$ , le calcul de la mesure  $S$  revient donc à calculer le nombre de ces structures claires de tailles supérieures à  $\eta$ . Par analogie, lorsque la mesure  $S$  est calculée pour la famille de fermetures  $(\varphi_\eta)_{\eta \geq 0}$ , ceci revient à calculer dans l'image  $I$ , le nombre de structures sombres de tailles supérieures à  $\eta$ . Par conséquent, nous

appelons un spectre de forme  $\mathcal{PS}(I, \eta)$  la fonction décrite par l'équation 3.32 et qui consiste à calculer le nombre de structures lumineuses de taille  $\eta$  dans une image  $I$ . Cette fonction indique la répartition de structures associées à chaque taille  $\eta$ . Par conséquent, le spectre de forme peut servir comme un descripteur de texture dans l'image. De manière générale, le spectre de forme  $\mathcal{PS}$  est souvent calculé sur la totalité de l'image tandis que, dans une même image, il est possible d'avoir plusieurs textures distinctes. Par conséquent, il est plus intéressant d'appliquer  $\mathcal{PS}(I_W, \eta)$  sur des fenêtres de taille  $W$  centrées sur les pixels de l'image. Nous appelons image d'énergie locale de taille  $\tau$  (seuil de la taille de structures claires) l'image formée des spectres de forme locaux de taille  $\tau$  de chaque pixel de l'image  $I$ .

$$\mathcal{PS}(I, \eta) = \begin{cases} S(\gamma_\eta(I)) - S(\gamma_{\eta+1}(I)) & \text{si } \eta \geq 0 \\ S(\varphi_\eta(I)) - S(\varphi_{\eta+1}(I)) & \text{si } \eta \leq -1 \end{cases} \quad (3.32)$$

#### d) Calcul du gradient morphologique de texture

Notons le gradient morphologique de texture  $\mathcal{G}^{tex}$  comme étant le résidu d'une dilatation morphologie et d'une érosion morphologique avec un élément structurant  $ES$ . Soit  $e_\tau$  l'image d'énergie locale de taille  $\tau$  calculée sur une couche de texture  $I_{tex}$ . Le gradient morphologique de texture est donc défini comme étant le supremum des gradients morphologiques des images d'énergie locales  $e_\tau$  lorsque  $\tau$  prend des valeurs dans l'ensemble  $\Pi = [-\eta, \eta]$ .

$$\mathcal{G}^{tex}(I_{tex})(p) = \bigvee_{\tau \in \Pi} [\delta_{ES}(e_\tau)(p) - \varepsilon_{ES}(e_\tau)(p)] \quad (3.33)$$

Le gradient de texture basé sur un filtre ASF consiste à appliquer l'analyse multi-échelle locale détaillée précédemment en utilisant comme image de marqueur dans le filtrage par nivellement, l'image simplifiée par le filtre ASF. Le processus de calcul du gradient  $\mathcal{G}_{ASF}$  est décrit par la figure 3.9. Dans nos expérimentations, l'élément structurant  $ES$  choisi pour les différentes opérations (l'ASF, la dilatation, l'érosion et l'ouverture morphologiques) est un disque de rayon adapté à chaque opération. Les fenêtres  $W$  centrées sur les pixels  $p$  de l'image ont été choisies empiriquement avec une même taille  $4 \times 4$  pixels. L'ordre choisi pour le filtre ASF est  $o = 6$  et l'ensemble  $\Pi = \{-2, -4, -6, 2, 4, 6\}$ .

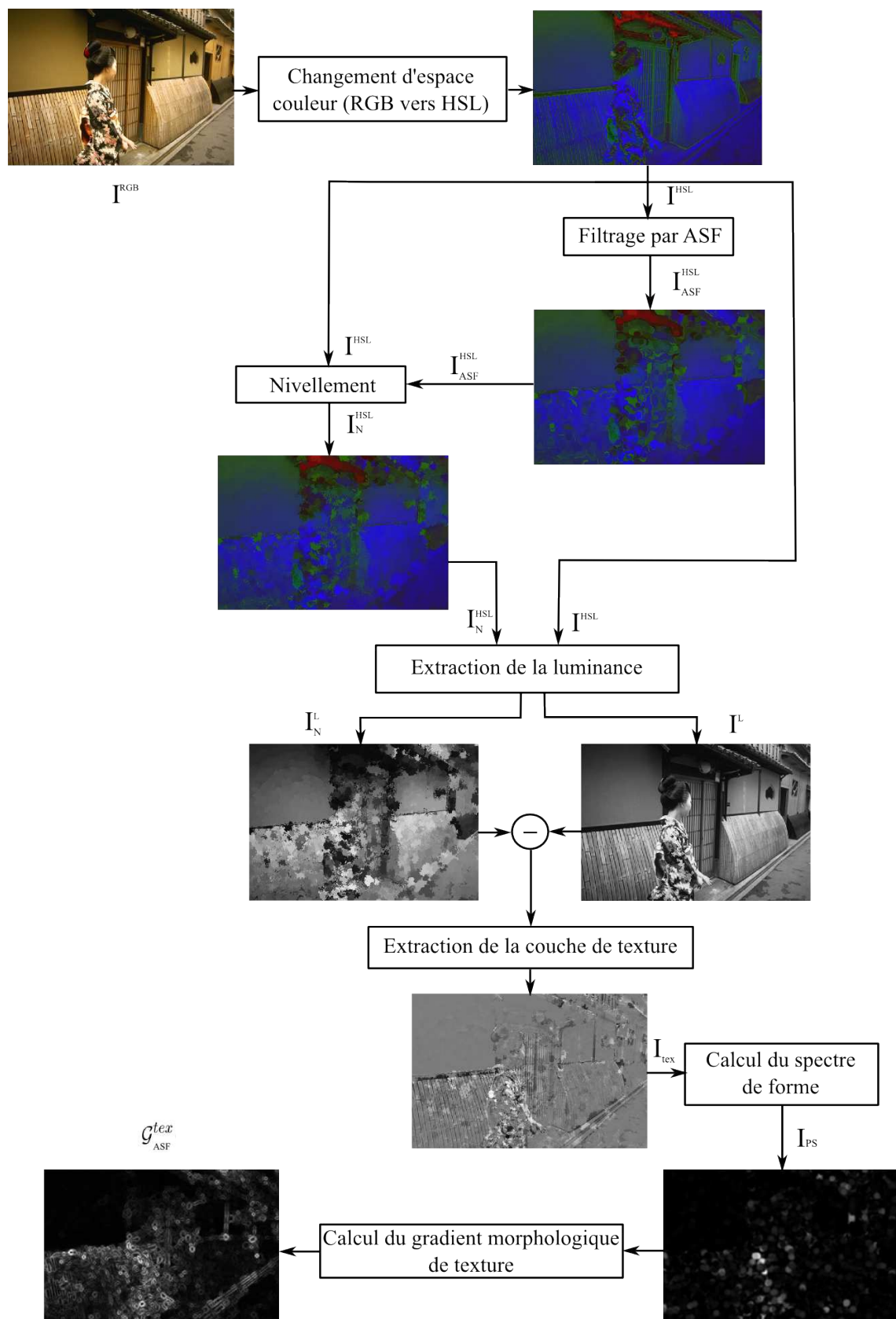


FIGURE 3.9 – Analyse multi-échelle locale pour le calcul du gradient morphologique de texture  $\mathcal{G}_{ASF}^{tex}$  basé sur le filtrage par ASF.

### 3.3/ UNE NOUVELLE MÉTHODE NON PARAMÉTRIQUE DE COMBINAISON COULEUR/TEXTURE

Maintenant que nous avons détaillé un certain nombre de descripteurs couleur et texture, nous allons dans cette section présenter une nouvelle méthode de combinaison couleur/texture basée sur l'analyse en composantes principales. La méthode proposée permet d'obtenir un gradient composite couleur/texture qui sera ensuite utilisé par l'algorithme de LPE en vue de segmenter l'image. Commençons donc par rappeler le principe de l'analyse en composantes principales (ACP).

#### 3.3.1/ L'ANALYSE EN COMPOSANTES PRINCIPALES

L'analyse en composantes principales (ACP) est une technique qui exploite l'aspect géométrique et les représentations graphiques d'un ensemble de données appelées observations, afin d'étudier sa variabilité et sa dispersion dans l'espace dans lequel il est représenté. Soit  $P$  un ensemble de données comportant  $m$  variables quantitatives ayant  $n$  unités (notées souvent individus), défini par l'équation 3.34.

$$P = \left[ \begin{pmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{i1} \\ \vdots \\ p_{n1} \end{pmatrix} \begin{pmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{i2} \\ \vdots \\ p_{n2} \end{pmatrix} \cdots \begin{pmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{ij} \\ \vdots \\ p_{nj} \end{pmatrix} \begin{pmatrix} p_{1m} \\ p_{2m} \\ \vdots \\ p_{im} \\ \vdots \\ p_{nm} \end{pmatrix} \right] \rightarrow \begin{bmatrix} (\vec{p}_1)^t \\ (\vec{p}_2)^t \\ \vdots \\ (\vec{p}_i)^t \\ \vdots \\ (\vec{p}_n)^t \end{bmatrix} \quad (3.34)$$

$$\downarrow$$

$$\left[ \vec{p}_{.1} \quad \vec{p}_{.2} \quad \cdots \quad \vec{p}_{.j} \quad \cdots \quad \vec{p}_{.m} \right]$$

En partant de son écriture matricielle, nous pouvons réécrire l'ensemble de données  $P$  sous deux formes vectorielles (cf. équation 3.34) : le vecteur vertical est composé d'un ensemble de points où chaque point, noté  $\vec{p}_i$ , est de dimension  $m$ . Le vecteur horizontal constitue la deuxième forme vectorielle et est composé d'un ensemble de points où chacun, noté  $\vec{p}_{.j}$ , est de dimension  $n$ . Notons que  $t$  désigne la fonction transposée. Les points  $\vec{p}_i$  et  $\vec{p}_{.j}$  peuvent être donnés par les équations suivantes :

$$\vec{p}_{.j} = \begin{pmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{pmatrix}_{1 \leq j \leq m} \quad \text{et} \quad \vec{p}_i = \begin{pmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{im} \end{pmatrix}_{1 \leq i \leq n} \quad (3.35)$$

où  $i$  et  $j$  indiquent respectivement l'individu et la variable.

Les représentations géométriques des unités permettent de déterminer les unités qui se ressemblent entre elles. Les représentations des variables permettent quant à elles d'étudier les structures de liaisons linéaires et de déterminer les variables qui sont corrélées [Duby 2006, Gergaud 2006]. Dans cette thèse, nous avons fait recours à l'ACP pour pouvoir extraire les axes principaux qui conservent le maximum d'information d'un

ensemble de données étudié [Gergaud 2006].

L'ACP est une technique basée sur l'algèbre linéaire. Elle consiste dans une première étape à calculer la matrice de covariance afin de déterminer les axes principaux de l'ensemble de données étudié. Ceci permet de décrire la dispersion des données par rapport à leurs axes principaux et de conclure sur les informations les plus pertinentes.

Dans ce qui suit, nous rappelons les définitions de la matrice de covariance  $COV_P$  de l'ensemble de données  $P$ . Calculer la matrice de covariance revient à calculer l'espérance  $\mathcal{E}$  du produit de l'ensemble de points  $P$  centré sur sa moyenne et du même vecteur résidu transposé, comme présenté par l'équation 3.36.

$$COV_P \equiv \mathcal{VAR}(P) \equiv \mathcal{E}((P - \mathcal{E}(P))(P - \mathcal{E}(P))^t) \quad (3.36)$$

En utilisant l'équation 3.34, nous pouvons redéfinir la matrice de covariance  $COV_P$  comme exprimé par l'équation 3.37.

$$COV_P \equiv \mathcal{VAR} \begin{bmatrix} (\vec{p}_{1.})^t \\ (\vec{p}_{2.})^t \\ \vdots \\ (\vec{p}_{i.})^t \\ \vdots \\ (\vec{p}_{n.})^t \end{bmatrix} \quad (3.37)$$

La matrice de covariance est formée par les variances/covariances de deux points  $\vec{p}_{.j}$  et  $\vec{p}_{.j'}$  [Gergaud 2006]. Elle peut être définie par l'écriture matricielle donnée par l'équation 3.38.

$$COV_P = \begin{pmatrix} var(\vec{p}_{1.}) & cov((\vec{p}_{1.})^t, \vec{p}_{2.}) & \cdots & cov((\vec{p}_{1.})^t, \vec{p}_{n.}) \\ cov((\vec{p}_{2.})^t, \vec{p}_{1.}) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ cov((\vec{p}_{n.})^t, \vec{p}_{1.}) & \cdots & \cdots & var(\vec{p}_{n.}) \end{pmatrix} \quad (3.38)$$

avec

$$cov((\vec{p}_{i.})^t, \vec{p}_{i.}) = var(\vec{p}_{i.}) \quad , \forall 1 \leq i \leq n \quad (3.39)$$

$$cov((\vec{p}_{i.})^t, \vec{p}_{i'.}) = \frac{1}{m} \sum_{j=1}^m (p_{ij} - \bar{p}_{i.})(p_{i'j} - \bar{p}_{i'.}) \quad , \forall 1 \leq i \leq n \quad (3.40)$$

$$\bar{p}_{i.} = \frac{1}{m} \sum_{j=1}^m p_{ij} \quad , \forall 1 \leq i \leq n \quad (3.41)$$

La matrice de covariance étant calculée, il s'agit ensuite d'extraire les valeurs propres et leurs vecteurs propres associés, correspondant aux axes principaux de l'ensemble de données étudié. Ceci permet ainsi d'obtenir une représentation géométrique donnant une idée de l'allure de ces données dans leur espace de représentation. Les valeurs propres indiquent la variance de l'ensemble de données (assimilé à un nuage de points de dimension  $m$ ) suivant les directions des axes principaux identifiés.

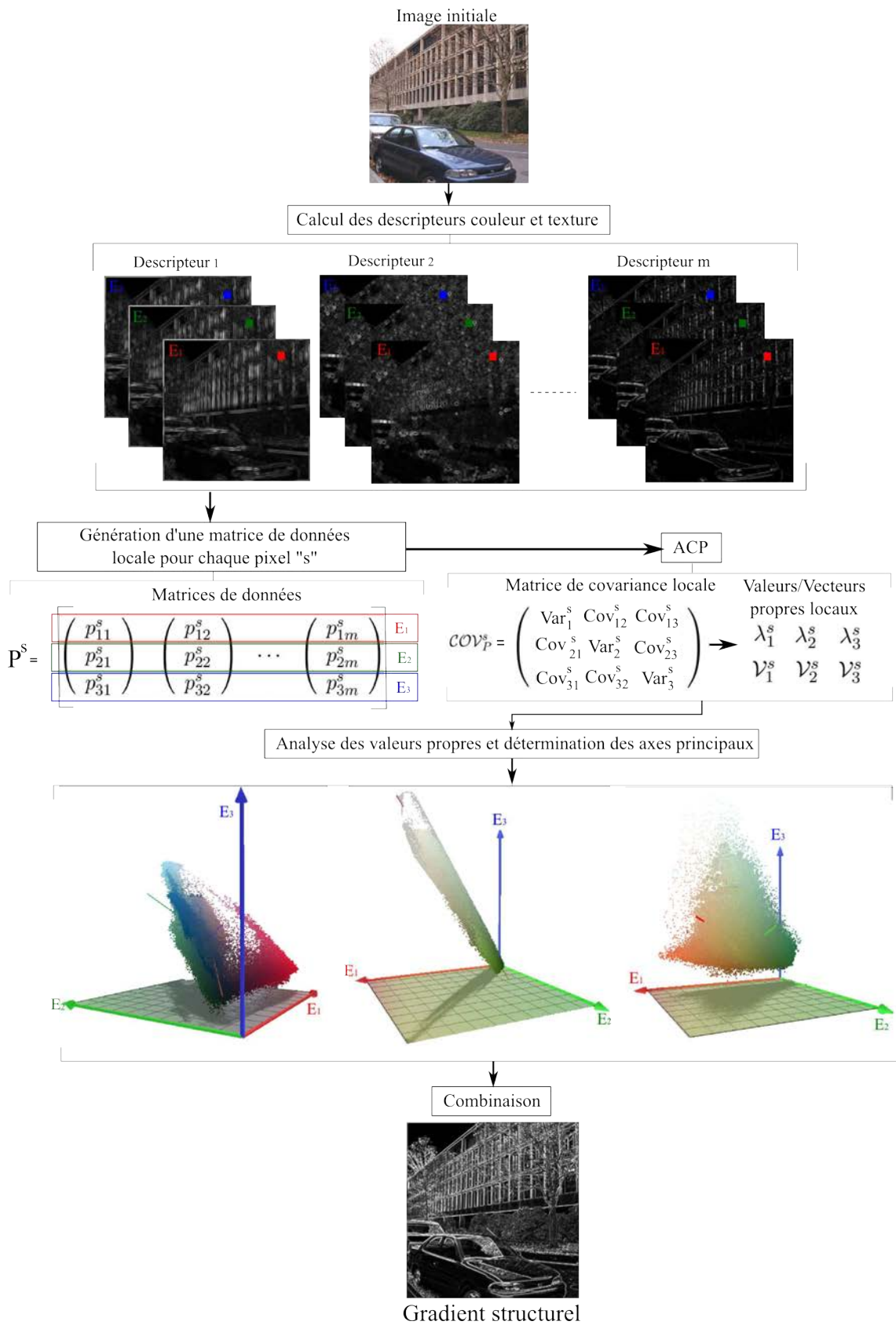


FIGURE 3.10 – Synopsis global de la combinaison couleur-texture basée sur l'ACP.



L'ACP est une technique souvent utilisée pour décorréler les données étudiées et réduire leur taille initiale. Dans le domaine de la vision par ordinateur, l'ACP a été utilisée dans de nombreuses applications, notamment la classification d'images [Maciejewski 2001, Fredembach 2004, Zeng 2011], la compression de l'information [Ho 2005, Mudrova, Du 2009], la reconnaissance des objets [Sirovich 1987, Turk 1991, Ilin 2010], etc. Dans cette thèse, nous utilisons l'ACP dans le but de proposer une nouvelle méthode de combinaison des informations de couleur et de texture présentes dans l'image. Plus précisément et comme précédemment mentionné, l'ACP permet d'extraire les axes qui conservent le maximum d'information. Nous utilisons donc cette technique afin d'identifier les informations (couleur et/ou texture) les plus pertinentes à combiner et ainsi obtenir un nouveau gradient structurel hybride permettant d'améliorer les résultats de segmentation par LPE. Dans ce qui suit, nous détaillons la stratégie globale de cette combinaison couleur/texture basée sur l'ACP, en présentant le calcul de la matrice de covariance et la méthode de combinaison des descripteurs calculés. Pour une meilleure compréhension, nous détaillons également (dans notre cas :  $n = 3$ ) le processus de calcul de la matrice de covariance sur les descripteurs couleur/texture de l'image à analyser ainsi que l'extraction des axes principaux des données obtenues (cf. figure 3.10).

### 3.3.2/ LA COMBINAISON COULEUR/TEXTURE PAR COMPARAISON DE VALEURS PROPRES

Commençons par définir l'ensemble de données sur lequel nous allons appliquer la combinaison couleur/texture proposée. Nous avons initialement calculé un ensemble de descripteurs couleur/texture de l'image à analyser.  $m$  désigne le nombre de descripteurs utilisés (les variables en considérant la terminologie de l'ACP).  $I$  désigne l'image initiale et  $s$  un pixel donné de  $I$ . Nous associons à  $s$  un ensemble de données  $P^s$  (un nuage de points 3D dans notre cas puisque le nombre d'unités  $n$  est égal à 3 pour chaque variable). En un pixel  $s$  de l'image, l'ensemble de données  $P^s$  est généré localement à partir des valeurs des  $m$  descripteurs de couleur et de texture. Les descripteurs utilisés sont des gradients couleur et des gradients texture calculés à partir de l'image initiale  $I$  dans un espace couleur donné  $E_1E_2E_3$ . Notons par  $\vec{g}_{.j}^s$  le  $j^{\text{ème}}$  gradient défini dans cet espace couleur, avec  $1 \leq j \leq m$ . Par analogie avec la notation de l'ACP, nous pouvons définir le nuage de points 3D selon les deux écritures matricielle (cf. équation 3.42) et vectorielle (cf. équation 3.43) suivantes :

$$P^s = \left[ \begin{pmatrix} g_{11}^s \\ g_{21}^s \\ g_{31}^s \end{pmatrix} \begin{pmatrix} g_{12}^s \\ g_{22}^s \\ g_{32}^s \end{pmatrix} \cdots \begin{pmatrix} g_{1m}^s \\ g_{2m}^s \\ g_{3m}^s \end{pmatrix} \right] \quad (3.42)$$

$$P^s = [\vec{g}_{.1}^s, \vec{g}_{.2}^s, \dots, \vec{g}_{.j}^s, \dots, \vec{g}_{.m}^s] \quad (3.43)$$

La deuxième étape consiste à appliquer l'ACP (cf. figure 3.10) en calculant la matrice de covariance du nuage de points  $P^s$ . La matrice de covariance de  $P^s$ , notée  $COV_P^s$ , est une matrice de taille  $3 \times 3$ , donnée par l'équation 3.44 :

$$COV_P^s = \begin{pmatrix} var(\vec{g}_{.1}^s) & cov((\vec{g}_{.1}^s)^t, \vec{g}_{.2}^s) & cov((\vec{g}_{.1}^s)^t, \vec{g}_{.3}^s) \\ cov((\vec{g}_{.2}^s)^t, \vec{g}_{.1}^s) & cov((\vec{g}_{.2}^s)^t, \vec{g}_{.2}^s) & cov((\vec{g}_{.2}^s)^t, \vec{g}_{.3}^s) \\ cov((\vec{g}_{.3}^s)^t, \vec{g}_{.1}^s) & cov((\vec{g}_{.3}^s)^t, \vec{g}_{.2}^s) & var(\vec{g}_{.3}^s) \end{pmatrix} \quad (3.44)$$

avec

$$\text{cov}((\vec{g}_{i.}^s)^t, \vec{g}_{i.}^s) = \text{var}(\vec{g}_{i.}^s) \quad , \forall 1 \leq i \leq 3 \quad (3.45)$$

$$\text{cov}((\vec{g}_{i.}^s)^t, \vec{g}_{i'.}^s) = \frac{1}{m} \sum_{j=1}^m (g_{ij}^s - \bar{g}_{i.}^s)(p_{i'j} - \bar{g}_{i'.}^s) \quad , \forall 1 \leq i \leq 3 \quad (3.46)$$

$$\bar{g}_{i.}^s = \frac{1}{m} \sum_{j=1}^m g_{ij}^s \quad , \forall 1 \leq i \leq 3 \quad (3.47)$$

Selon les règles de l'algèbre linéaire, il existe deux matrices notés  $L^s$  et  $D^s$ , telles que :

$$(L^s)^{-1} COV_p^s L^s = D^s \quad (3.48)$$

La matrice de covariance étant de taille  $3 \times 3$ , il en résulte trois valeurs propres et trois vecteurs propres associés. Nous désignons par  $\lambda_1^s$ ,  $\lambda_2^s$  et  $\lambda_3^s$ , les valeurs propres de  $COV_p^s$  telles que  $\lambda_3^s \leq \lambda_2^s \leq \lambda_1^s$ , et par  $\vec{\mathcal{V}}_1^s$ ,  $\vec{\mathcal{V}}_2^s$ ,  $\vec{\mathcal{V}}_3^s$  les vecteurs propres associés. Les matrices  $D^s$  et  $L^s$  sont alors définies par les deux équations données ci-après :

$$D^s = \begin{pmatrix} \lambda_1^s & 0 & 0 \\ 0 & \lambda_2^s & 0 \\ 0 & 0 & \lambda_3^s \end{pmatrix} \quad (3.49)$$

$$L^s = [\vec{\mathcal{V}}_1^s \quad \vec{\mathcal{V}}_2^s \quad \vec{\mathcal{V}}_3^s] \quad (3.50)$$

Pour déterminer les axes principaux qui mettent en évidence le maximum d'informations et décider ensuite quelles sont les informations à combiner, nous procédons à une comparaison des valeurs propres. Notons par  $\hat{G}$  le gradient structural obtenu après la combinaison et  $\hat{G}(s)$  sa valeur estimée en  $s$ . Le processus de comparaison distingue trois cas décrits ci-après :

**Cas 1**  $\lambda_1^s \gg \lambda_2^s$  (à titre informatif si  $\frac{\lambda_1^s}{\lambda_2^s} \geq 10^2$ ) : Il existe un seul axe principal dont la direction est celle du vecteur propre  $\vec{\mathcal{V}}_1^s$ , autour duquel le nuage de points 3D est concentré (cf. figure 3.11(a)). Dans ce cas, nous choisissons la seule composante principale qui conserve le maximum d'informations et dont la direction est donnée par le vecteur propre  $\vec{\mathcal{V}}_1^s$ . La valeur du gradient structural en un pixel  $s$  est donnée par la formulation mathématique suivante :

$$\hat{G}(s) = \text{argmax}_{j \in \{1, \dots, m\}} \{g_{1j}\}$$

**Cas 2** :  $\lambda_3^s \ll \lambda_2^s$  (à titre informatif si  $\frac{\lambda_2^s}{\lambda_3^s} \geq 10$ ) : Dans ce cas,  $\lambda_1^s$  et  $\lambda_2^s$  ont le même ordre de grandeur. Le nuage de points 3D s'apparente à un plan formé par les deux vecteurs propres  $\vec{\mathcal{V}}_1^s$  et  $\vec{\mathcal{V}}_2^s$  (associés aux valeurs propres  $\lambda_2^s$  et  $\lambda_1^s$ ) (cf. figure 3.11(b)). Le vecteur propre  $\vec{\mathcal{V}}_3^s$  constitue la normale du plan formé par les vecteurs propres  $\vec{\mathcal{V}}_1^s$  et  $\vec{\mathcal{V}}_2^s$ . Par conséquent, les composantes principales qui conservent le maximum d'informations sont celles qui forment le plan  $(\vec{\mathcal{V}}_1^s, \vec{\mathcal{V}}_2^s)$ . La valeur du gradient structural en un pixel  $s$  est alors donnée par la formulation mathématique suivante :

$$\hat{G}(s) = \operatorname{argmax}_{j \in \{1, \dots, m\}} \left\{ \frac{1}{2} (g_{1j}^s + g_{2j}^s) \right\}$$

**Cas 3 :**  $\lambda_1^s \approx \lambda_2^s \approx \lambda_3^s$  : Dans ce dernier cas, le nuage de points 3D est dispersé dans toutes les directions (cf. figure 3.11(c)) et aucune information n'est alors privilégiée. Nous combinons donc toutes les composantes équitablement. La valeur du gradient structurel en un pixel  $s$  est donnée par la formulation mathématique suivante :

$$\hat{G}(s) = \operatorname{argmax}_{j \in \{1, \dots, m\}} \left\{ \frac{1}{3} (g_{1j}^s + g_{2j}^s + g_{3j}^s) \right\}$$

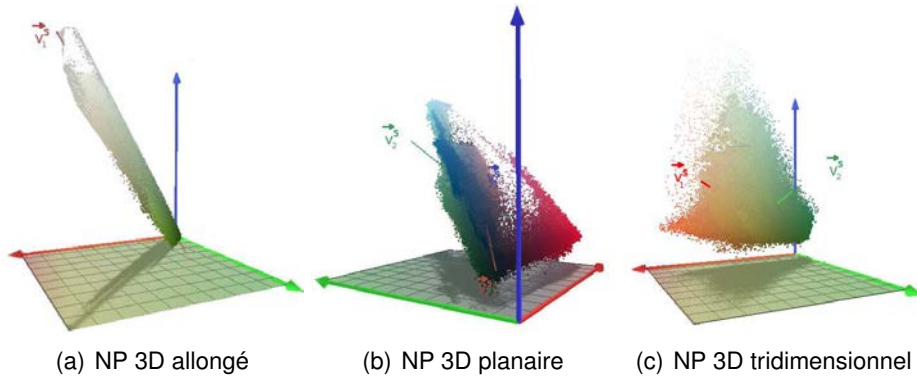


FIGURE 3.11 – Exemple de trois dispersions géométriques possibles d'un nuage de points (NP) 3D.

L'algorithme 6 donné ci-après reprend les différentes étapes du calcul du gradient structurel, basé sur la combinaison couleur/texture non paramétrique proposée.

Après avoir détaillé notre méthode de combinaison couleur/texture non paramétrique, basée sur l'analyse des composantes principales et permettant d'obtenir un gradient structurel en niveau de gris, nous allons rappeler l'algorithme de segmentation par Ligne de Partage des Eaux qui fait appel à ce dernier (gradient structurel).

**Algorithme 6:** Méthode de combinaison couleur/texture non paramétrique proposée**Entrées :**  $I$  une image couleur.Calculer les  $m$  différents descripteurs sur l'image  $I$  dans l'espace couleur  $E_1E_2E_3$ .**Répéter**

1. Générer l'ensemble des données

$$P^s = [g^s_{.1}, g^s_{.2}, \dots, g^s_{.j}, \dots, g^s_{.m}]$$

2. Calculer la matrice de covariance locale

$$COV_P^s = \begin{pmatrix} var(g^s_{.1}) & cov((g^s_{.1})^t, g^s_{.2}) & cov((g^s_{.1})^t, g^s_{.3}) \\ cov((g^s_{.2})^t, g^s_{.1}) & cov((g^s_{.2})^t, g^s_{.2}) & cov((g^s_{.2})^t, g^s_{.3}) \\ cov((g^s_{.3})^t, g^s_{.1}) & cov((g^s_{.3})^t, g^s_{.2}) & var(g^s_{.3}) \end{pmatrix}$$

3. Extraire les valeurs propres
- $\lambda_1^s$
- ,
- $\lambda_2^s$
- et
- $\lambda_3^s$
- et leurs vecteurs propres associés
- $\vec{\mathcal{V}}_1^s$
- ,
- $\vec{\mathcal{V}}_2^s$
- et
- $\vec{\mathcal{V}}_3^s$
- .

- 4.
- Suivant**
- la valeur de
- $\{\lambda_i^s\}_{1 \leq i \leq 3}$
- faire**

**Cas où**  $\lambda_1^s \gg \lambda_2^s$ 

$$\hat{G}(s) = \operatorname{argmax}_{j \in \{1, \dots, m\}} \{g^s_{3j}\}$$

**fin****Cas où**  $\lambda_3^s \ll \lambda_2^s$ 

$$\hat{G}(s) = \operatorname{argmax}_{j \in \{1, \dots, m\}} \{\frac{1}{2}(g^s_{1j} + g^s_{2j})\}$$

**fin****Cas où**  $\lambda_1^s \simeq \lambda_2^s \simeq \lambda_3^s$ 

$$\hat{G}(s) = \operatorname{argmax}_{j \in \{1, \dots, m\}} \{\frac{1}{3}(g^s_{1j} + g^s_{2j} + g^s_{3j})\}$$

**fin****fin****Jusqu'à** la fin de pixel  $s$  dans l'image  $I$ ;**Sorties :** Gradient structurel  $\hat{G}$

### 3.4/ LA MÉTHODE DE LIGNE DE PARTAGE DES EAUX

La technique de Ligne de Partage des Eaux (LPE), issue de la morphologie mathématique, est l'une des méthodes de segmentation morphologique les plus communément utilisées dans la littérature. Cette technique consiste à considérer l'image à segmenter comme un relief (surface topographique) où les niveaux de gris de l'image constituent son altitude. En effet, la méthode de LPE est à la base une notion interprétée par analogie géographique comme le complémentaire des bassins versants [Angelini 2007]. ANGELINI et al. définissent un bassin versant comme étant la zone associée à un minimum régional telle qu'une goutte d'eau tombant dans cette zone suivant la ligne de plus grande pente s'arrête dans ce minimum [Angelini 2007]. L'illustration de ce phénomène est donnée par la figure 3.12.

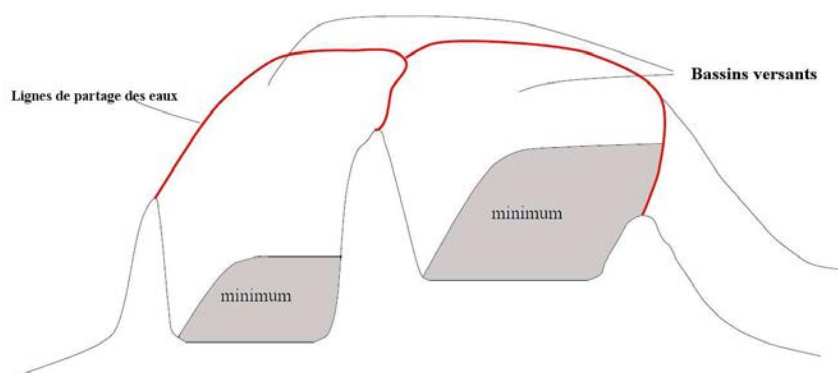


FIGURE 3.12 – Bassins versants et ligne de partage des eaux [Vincent 1991].

Le paradigme d'immersion introduit par VINCENT et SOILLE [Vincent 1991] est l'un des algorithmes les plus connus pour la construction de LPE. Ce paradigme consiste à remplir les bassins versants, obtenus par les minimums régionaux, pour arriver à leurs limites qui constituent les LPE. Des travaux plus récents permettent de donner des définitions plus rigoureuses de la méthode de LPE [Schmitt 1994].

L'avantage de la méthode de LPE en tant que méthode de segmentation est qu'elle permet de fournir des régions bornées par des contours fermés. Dans le cas usuel, les minimums régionaux qui définissent les bassins versants sont souvent déterminés par le gradient de l'image représentant les contours. En appliquant la méthode de LPE sur le gradient de l'image, les bassins versants devraient théoriquement correspondre à des régions homogènes de cette image. Après la détermination des contours de ces bassins, qui forment les minima locaux dans l'image, nous obtenons une carte de régions. La méthode de LPE la plus répandue utilise en entrée deux images : une image de potentiel (gradient) et une image de germes [Shafarenko 1997].

Dans la pratique, cette transformation produit une importante sur-segmentation due au bruit ou aux irrégularités locales dans l'image de potentiel. En effet, ces irrégularités influencent le nombre de germes (généralement le nombre de minima locaux) qui devient souvent très important et engendre la sur-segmentation en question. Plusieurs alternatives ont été proposées dans la littérature pour remédier à cet inconvénient. Une

première alternative consiste à utiliser une méthode de Ligne de Partage des Eaux hiérarchique non paramétrique, connue également sous le nom d'algorithme de Waterfall [Beucher 1994]. Pour cette raison, différents gradients ont été définis [Angulo 2003, Lezoray 2008]. L'inconvénient majeur de cette solution est qu'elle nécessite la définition d'un niveau optimal de la hiérarchie (autrement dit celui qui offre la meilleure segmentation) et que le temps de traitement est important. Une deuxième alternative consiste à agir sur l'image de germes. Dans le cadre de cette alternative, nous avons récemment proposé [Cohen 2010a] une solution permettant de réduire le nombre de minima locaux dans l'image, en fonction de deux paramètres fixés expérimentalement. Le principe général de cette solution est présenté ci-après.

### SÉLECTION AUTOMATIQUE DES GERMES

L'opération de sélection de germes proposée dans les travaux de COHEN et al. [Cohen 2010a], permet de réduire le nombre de minima dans l'image de germes en fonction de deux paramètres  $\nu_1$  et  $\nu_2$ . Ces derniers définissent le pourcentage de minima locaux à conserver lors de l'étape de croissance de régions. Cette sélection permet d'agir sur le nombre de germes tout en conservant l'entropie de leur répartition dans l'image.

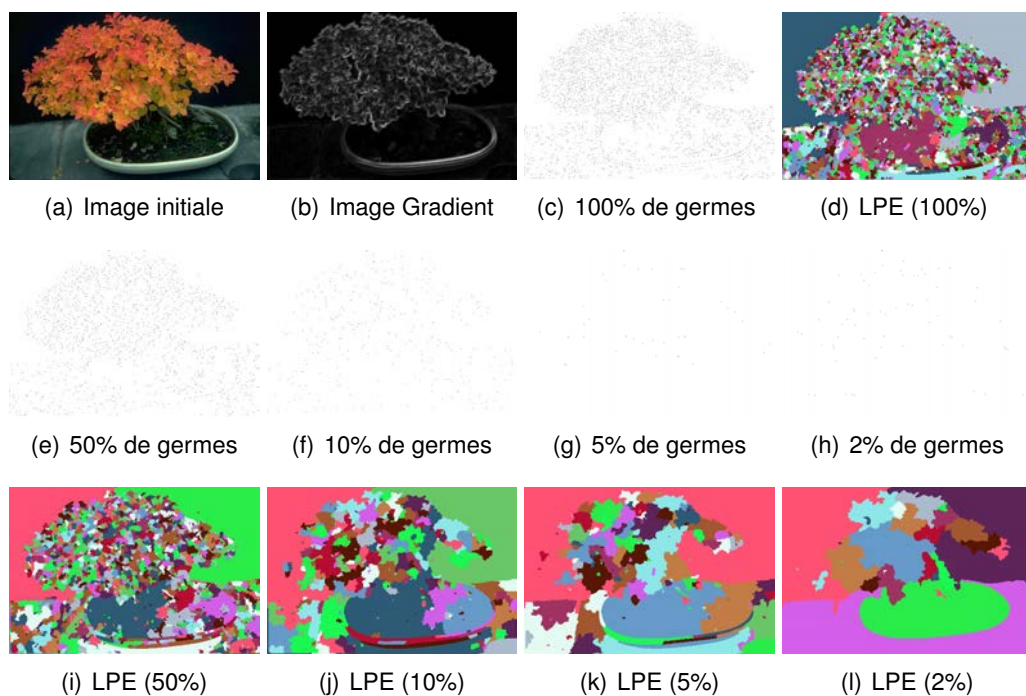


FIGURE 3.13 – Exemples de réductions de germes avec le résultat de segmentation correspondant.

La réduction du nombre de minima locaux garantie une meilleure segmentation en éliminant les petites régions insignifiantes. Le processus de sélection commence par un premier parcours de l'image de germes pour sauvegarder les minima locaux sans tenir compte de leur position dans l'image. En effet, en fonction des paramètres d'entrée  $\nu_1$  et  $\nu_2$ , on effectue, régulièrement, la sélection de  $\nu_1$  germes de chaque  $\nu_2$ -uplet. Les germes sélectionnés sont stockés dans un deuxième tableau. Un deuxième parcours de



l'image de germes est ensuite réalisée de sorte à éliminer les marqueurs (les germes) non considérés en vérifiant leur absence dans le nouveau tableau. La figure 3.13 illustre le résultat de la réduction de germes ainsi que son effet sur la segmentation d'une image de la base de BERKELEY.

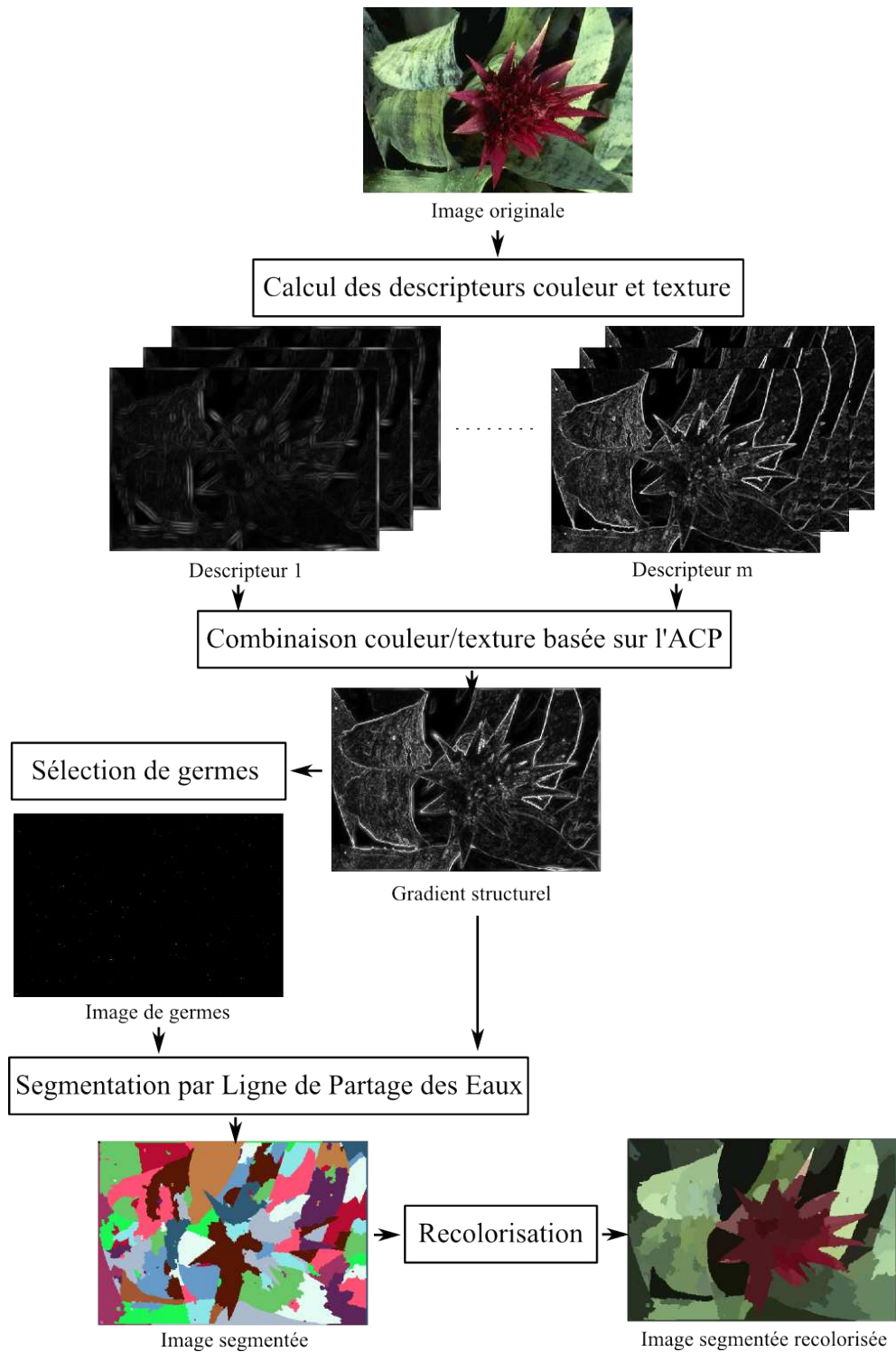


FIGURE 3.14 – Synopsis global de la méthode de segmentation proposée.

La figure 3.14 illustre l'approche globale proposée dans le cadre de cette thèse. Dans la suite de ce chapitre, nous effectuons une étude qualitative et quantitative des résultats de segmentation basée sur la nouvelle combinaison couleur/texture. Nous situons également les performances de cette méthode par rapport aux principales méthodes de segmentation de la littérature.

### 3.5/ RÉSULTATS EXPÉRIMENTAUX

Dans cette section, nous analysons la performance de la combinaison couleur/texture proposée, en évaluant la qualité des images segmentées par l'algorithme de LPE. Cette performance est ensuite comparée aux performances des autres méthodes de segmentation de la littérature, connues pour leurs très bons résultats. Nous rappelons que l'évaluation de ces méthodes a été présentée dans le chapitre 2. Tout comme les combinaisons fixe, adaptative et supremum, la performance de la combinaison proposée est mesurée en fonction des couples de gradients couleur/texture utilisés. Nous rappelons au lecteur que nous reprenons les différents gradients couleur et texture introduits dans ce chapitre, à savoir les gradients de DI-ZENZO et morphologique couleur pour les descripteurs couleur et les gradients basés sur les filtres de Gabor, les attributs  $H_1$ ,  $H_2$  et  $H_5$  d'HARALICK et l'ASF pour les descripteurs de texture. Notons aussi que l'élément structurant utilisé pour le calcul du gradient texture basé sur l'ASF est de forme circulaire. Nous précisons que l'image de germes utilisée est générée à partir du gradient structurel obtenu, suite à une réduction des minima locaux, avec  $v_1 = 1$  et  $v_2 = 50$  (un pourcentage de 2%) des minima locaux.

Nous rappelons également au lecteur que nous avons sélectionné deux ensembles de métriques d'évaluation. Le premier concerne les métriques d'évaluation supervisée avec la mesure de Rand probabiliste  $\mathcal{PR}$ , la mesure de JACCARD  $\mathcal{J}$ , la mesure de FOWLKES et MALLOWS  $\mathcal{F}$ , la mesure de Bipartite Graph Matching  $\mathcal{BGM}$ , la mesure d'information mutuelle normalisée  $\mathcal{NMI}$ , la mesure de variation d'information  $\mathcal{VI}$ , l'erreur locale de cohérence  $\mathcal{LCE}$ , l'erreur globale de cohérence  $\mathcal{GCE}$  et la mesure de VINET  $\mathcal{V}$ . Le deuxième ensemble de mesures utilisées concerne les métriques d'évaluation non-supervisée avec l'erreur quadratique moyenne  $\mathcal{MSE}$ , la mesure de ZEBOUJ  $\mathcal{Z}$  et la mesure de BORSOTTI  $\mathcal{B}$ . Les métriques  $\mathcal{PR}$ ,  $\mathcal{J}$ ,  $\mathcal{F}$ ,  $\mathcal{V}$ ,  $\mathcal{NMI}$  et  $\mathcal{Z}$  indiquent une meilleure performance lorsqu'elles tendent vers 1 alors que les autres métriques indiquent une meilleure performance lorsque leurs valeurs sont faibles.

Le tableau 3.1 présente l'évaluation moyenne, sur 200 images de la base de BERKELEY, des images segmentées par l'algorithme de LPE, en utilisant la combinaison couleur/texture proposée. La performance de l'approche proposée, présentée en fonction des couples gradients couleur/texture, est mesurée à l'aide de métriques d'évaluation présentées dans le chapitre 2. À partir du tableau 3.1, nous constatons que pour la combinaison proposée, le couple de gradients couleur/texture H2M (correspondant au couple : gradient texture basé sur le 2<sup>ème</sup> paramètre d'HARALICK et le gradient morphologique couleur) obtient la première position avec la majorité des métriques d'évaluation utilisées (9/12 fois en première position).

		Combinaison basée sur l'Analyse en Composantes Principales									
		AM	AZ	GM	GZ	H1M	H1Z	H2M	H2Z	H5M	H5Z
Critères d'évaluation	<i>PR</i>	<b>0.79</b>	<b>0.79</b>	0.78	<b>0.79</b>	<b>0.79</b>	<b>0.79</b>	<b>0.79</b>	<b>0.79</b>	<b>0.79</b>	<b>0.79</b>
	<i>J</i>	0.15	0.13	0.14	0.11	0.11	0.10	<b>0.17</b>	0.16	0.15	0.12
	<i>F</i>	0.31	0.28	0.28	0.26	0.26	0.24	<b>0.32</b>	0.31	0.31	0.27
	<i>BGM</i>	0.10	0.09	0.10	0.08	0.08	0.07	<b>0.12</b>	<b>0.12</b>	0.11	0.09
	<i>NMI</i>	0.12	<b>0.13</b>	0.12	<b>0.13</b>	0.12	<b>0.13</b>	<b>0.12</b>	<b>0.13</b>	0.12	<b>0.13</b>
	<i>VI</i>	4.87	5.21	4.97	5.29	5.37	5.55	<b>4.79</b>	5.19	4.84	5.26
	<i>LCE</i>	<b>0.23</b>	<b>0.23</b>	0.24	<b>0.23</b>	0.25	0.24	<b>0.23</b>	0.24	<b>0.2</b>	<b>0.23</b>
	<i>GCE</i>	0.35	0.32	0.36	0.32	0.34	<b>0.31</b>	<b>0.35</b>	0.34	0.35	<b>0.31</b>
	<i>V</i>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>	<b>0.95</b>	<b>0.96</b>	0.95
	<i>MSE</i> (10 <sup>2</sup> )	4.07	3.47	4.15	<b>3.41</b>	4.2	3.56	<b>4.13</b>	3.5	4.09	3.28
	<i>Z</i>	0.46	0.41	0.41	0.38	0.38	0.22	<b>0.49</b>	0.42	0.44	0.39
	<i>B</i>	<b>0.12</b>	0.14	<b>0.12</b>	0.13	0.14	0.14	<b>0.12</b>	0.14	<b>0.12</b>	0.13

TABLE 3.1 – Performances moyennes (calculées sur 200 images de la base de BERKELEY) avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture de la méthode de segmentation par la combinaison couleur/texture proposée.

Dans ce qui suit, nous comparons les résultats de segmentation obtenus avec l'approche de combinaison proposée (sur 200 images de la base de BERKELEY) par rapport aux résultats des méthodes de segmentation de la littérature introduites dans le chapitre 1 et évaluées dans le chapitre 2. Les résultats de segmentation obtenus avec les paramètres optimaux sont donc synthétisés dans le tableau 3.2. Contrairement à la comparaison effectuée dans le chapitre 2, pour laquelle nous considérons que l'évaluation moyenne ne donne pas nécessairement une évaluation précise et décisive de la performance d'une méthode de segmentation par rapport à d'autres, nous avons rajouté dans ce chapitre des propriétés statistiques pour chaque métrique d'évaluation. Ceci permettra de mieux contrôler la performance de chaque méthode de segmentation. Les propriétés statistiques considérées sont :

- **Moy** indique la valeur moyenne de la métrique utilisée sur la totalité des images de la base de test.
- **Min** indique la valeur minimale de la métrique utilisée sur la totalité des images de la base de test.
- **Max** indique la valeur maximale de la métrique utilisée sur la totalité des images de la base de test.
- **Med** indique la valeur médiane de la métrique utilisée sur la totalité des images de la base de test.
- **ET** indique l'écart type de la métrique utilisée sur la totalité des images de la base de test.
- **Sk** indique l'asymétrie de la métrique utilisée (*skewness* en anglais) sur la totalité des images de la base de test. En théorie de probabilité et statistique, il s'agit du

premier paramètre de forme et le moment d'ordre 3 d'une variable centrée réduite. L'asymétrie de la métrique est positive si la distribution de la métrique s'étale plus vers les valeurs élevées de la métrique (distribution décalée à gauche de sa moyenne) et elle est négative si la distribution s'étale plus vers les valeurs basses de la métrique (distribution décalée à droite de sa moyenne). L'asymétrie est nulle si la distribution de la métrique est parfaitement symétrique.

- **Ku** indique l'aplatissement de la distribution de la métrique utilisée sur la totalité des images de la base de test. En théorie de probabilité et statistique, il s'agit du deuxième paramètre de forme et le moment d'ordre 4 d'une variable centrée réduite. Si la valeur de l'aplatissement est supérieur à 3, cela veut dire que la distribution de la métrique est plus tranchante qu'une distribution normale, avec des valeurs concentrées autour de la moyenne et des queues plus épaisses. En d'autres termes, cela signifie une forte probabilité pour les valeurs extrêmes. Cependant, si la valeur de l'aplatissement est inférieure à 3 ; cela veut dire que la distribution de la métrique est plus plate qu'une distribution normale avec un pic plus large. La probabilité d'avoir des valeurs extrêmes est plus faible que pour une distribution normale, et les valeurs sont réparties de façon plus étalée autour de la moyenne. Une valeur d'aplatissement égale à 3 caractérise une distribution normale.

Pour une meilleure visualisation des résultats, nous tenons à informer le lecteur que les valeurs en rouge (respectivement en bleu) correspondent aux premiers (respectivement deuxièmes) meilleurs résultats. Les cellules en vert correspondent quant à elles à la meilleure méthode de segmentation selon la métrique d'évaluation considérée.

D'après le tableau 3.2, nous pouvons constater que la méthode de segmentation CTB basée sur la combinaison des informations de couleur, de texture et de luminance se positionne comme la meilleure méthode de segmentation selon la plupart des métriques d'évaluation. En effet, pour chaque métrique d'évaluation utilisée, la meilleure méthode de segmentation est déterminée en examinant les meilleures valeurs des différentes propriétés statistiques utilisées, à savoir la valeur moyenne  $Moy$ , la valeur minimale  $Min$ , la valeur maximale  $Max$ , la valeur médiane  $Med$ , l'écart type  $ET$ , l'asymétrie  $Sk$  et l'aplatissement de distribution  $Ku$ . Prenons l'exemple de la mesure de BORSOTTI  $\mathcal{B}$  : la segmentation par CTB donne la deuxième meilleure valeur moyenne  $Moy = 0.06$ , d'asymétrie  $Sk = 0.73$  et d'aplatissement  $Ku = 3.33$  ; et la première meilleure valeur minimale  $Min = 0.01$ , maximale  $Max = 0.15$ , médiane  $Med = 0.05$  et écart type  $ET = 0.03$ . Par conséquent, si nous considérons la métrique  $\mathcal{B}$ , la méthode de segmentation CTB, comptant quatre premières meilleures valeurs et trois deuxièmes meilleures valeurs sur sept propriétés statistiques, se positionne en première position par rapport aux autres méthodes de segmentation. D'autre part, la méthode de segmentation CTB est favorisée par cinq métriques  $\mathcal{J}$ ,  $\mathcal{F}$ ,  $\mathcal{VI}$ ,  $\mathcal{GVE}$  et  $\mathcal{B}$ , et se positionne donc comme la meilleure méthode de segmentation (5 fois sur 11 en première position). La deuxième meilleure méthode de segmentation est le MeanShift (MS) (2 fois sur 11 en première position). Nous pouvons également constater que si la combinaison proposée (basée sur l'ACP) n'occupe pas la première ou la deuxième position (couleur rouge ou bleue dans le tableau 3.2), elle se situe souvent en bonne position (oscille entre 2<sup>ème</sup> et 6<sup>ème</sup> position) et l'écart numérique d'évaluation par rapport aux méthodes qui la précèdent est remarquablement faible.



Critères d'évaluation		MS	JSEG	FH	FCR	NCUTS	CTM	CV	CTB	CF	CA	CS	C-ACP
Évaluation supervisée													
NMI	Moy	0.06	0.04	0.11	0.08	0.10	0.06	0.08	0.09	0.08	0.13	0.13	0.12
	Min	0	0.01	0.01	0	0.03	0.01	0.01	0.01	0	0.03	0.02	0.04
	Max	0.12	0.11	0.21	0.14	0.15	0.12	0.14	0.15	0.16	0.23	0.23	0.21
	Med	0.06	0.04	0.12	0.08	0.10	0.07	0.08	0.09	0.08	0.13	0.13	0.12
	ET	0.03	0.02	0.04	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.03
	Sk	-0.14	0.48	-0.40	-0.59	-0.57	-0.33	-0.20	-0.43	-0.21	-0.24	-0.31	-0.18
	Ku	2.26	3.21	2.95	2.90	2.75	2.52	2.63	2.84	2.58	3.13	3.13	2.94
VI	Moy	4.15	6.06	4	3.86	4.05	4.05	5.21	3.54	3.96	4.58	4.52	4.79
	Min	1.74	4.75	1.96	1.65	2.62	1.85	2.09	1.43	1.58	2.62	2.61	2.93
	Max	7.09	7.34	6.31	6.62	5.46	6.63	8.25	6.81	5.52	6.11	5.95	6.41
	Med	4.03	6.08	3.92	3.91	4.08	4.08	5.23	3.47	4	4.61	4.56	4.78
	ET	1.15	0.45	0.86	0.88	0.50	0.90	1.06	0.92	0.78	0.74	0.71	0.69
	Sk	0.39	-0.08	0.03	0.17	-0.09	0.29	0.15	0.30	-0.39	-0.36	-0.23	-0.11
	Ku	2.69	3.03	2.60	3.03	2.79	3.22	2.96	2.92	2.93	2.53	2.48	2.41
LCE	Moy	0.29	0.56	0.26	0.28	0.22	0.26	0.31	0.23	0.18	0.22	0.22	0.23
	Min	0	0.05	0.05	0	0.06	0.03	0.08	0.03	0.01	0.07	0.06	0.07
	Max	0.62	0.67	0.43	0.60	0.35	0.54	0.55	0.59	0.37	0.34	0.33	0.37
	Med	0.27	0.57	0.25	0.28	0.22	0.25	0.31	0.23	0.19	0.23	0.23	0.23
	ET	0.15	0.08	0.08	0.11	0.05	0.10	0.08	0.10	0.07	0.06	0.06	0.05
	Sk	0.35	-2.12	-0.04	-0.08	0.06	0.18	0.19	0.51	-0.10	-0.14	-0.32	-0.31
	Ku	2.50	12.06	2.49	3.33	2.89	2.81	2.90	3.25	2.96	2.46	2.53	2.91
GCE	Moy	0.32	0.63	0.37	0.34	0.32	0.31	0.41	0.27	0.29	0.33	0.33	0.35
	Min	0.01	0.07	0.07	0	0.12	0.05	0.10	0.03	0.02	0.09	0.08	0.08
	Max	0.70	0.75	0.60	0.65	0.51	0.67	0.64	0.65	0.55	0.51	0.54	0.54
	Med	0.31	0.64	0.37	0.34	0.33	0.31	0.41	0.27	0.30	0.34	0.33	0.36
	ET	0.15	0.08	0.11	0.12	0.08	0.12	0.10	0.11	0.11	0.10	0.10	0.09
	Sk	0.26	-2.24	-0.24	-0.27	-0.10	0.09	-0.15	0.36	-0.14	-0.24	-0.24	-0.31
	Ku	2.56	12.92	2.52	3.33	2.74	2.66	2.89	3.03	2.48	2.22	2.50	2.76
Suite à la page suivante													



		MS	JSEG	FH	FCR	NcUTS	CTM	CV	CTB	CF	CA	CS	C-ACP
Critères d'évaluation		Évaluation non supervisée											
$MSE$	<i>Moy</i>	285.3	252.4	380.2	288.9	506.4	478.6	429.8	392.7	642.5	353.6	331.6	412.9
	<i>Min</i>	75.6	50.2	97.4	102	126.6	90.7	120.6	119.2	138.4	104.8	80.3	100
	<i>Max</i>	1517.1	856.4	1103.4	633.2	1459.1	1564.6	1539	1669.2	1965.1	906.8	825	1215.9
	<i>Med</i>	241.1	228.7	367	272.6	455.9	451.6	384.4	373.4	565.9	333.1	304.5	362.2
	<i>ET</i>	200.6	135.3	158.0	110.2	246.7	233.2	211.5	177.9	338.5	151.4	142.8	201.4
	<i>Sk</i>	3.2	1.3	1.2	0.8	1.1	1.2	2.2	2.4	1.2	1	0.9	1.1
	<i>Ku</i>	16.7	5.7	5.8	3.6	4.7	5.4	10.6	15.7	4.8	4.7	3.9	4.5
$\mathcal{Z}$	<i>Moy</i>	0.61	0.12	0.52	0.28	0.33	0.29	0.28	0.35	0.27	0.43	0.45	0.49
	<i>Min</i>	0.15	0	0	0	0.03	0.02	0	0.01	0	0.06	0.10	0.19
	<i>Max</i>	0.84	0.45	0.75	0.71	0.67	0.66	0.59	0.72	0.66	0.75	0.76	0.76
	<i>Med</i>	0.63	0.09	0.53	0.28	0.33	0.28	0.28	0.35	0.26	0.42	0.45	0.48
	<i>ET</i>	0.14	0.10	0.11	0.16	0.14	0.15	0.11	0.15	0.17	0.15	0.14	0.12
	<i>Sk</i>	-0.56	0.88	-1.25	0.14	0.01	0.14	0.31	0.03	0.32	-0.08	0.04	-0.03
	<i>Ku</i>	2.99	2.98	7.66	2.22	2.20	2.06	2.65	2.40	2.24	2.39	2.47	2.48
$\mathcal{B}$	<i>Moy</i>	0.07	0.09	0.13	0.05	0.18	0.07	0.20	0.06	0.11	0.10	0.10	0.12
	<i>Min</i>	0.01	0.02	0.02	0.01	0.04	0.01	0.04	0.01	0.01	0.02	0.02	0.02
	<i>Max</i>	0.18	0.27	0.31	0.16	0.62	0.18	0.76	0.15	0.40	0.39	0.34	0.46
	<i>Med</i>	0.06	0.08	0.12	0.05	0.17	0.06	0.19	0.05	0.10	0.10	0.09	0.11
	<i>ET</i>	0.03	0.05	0.06	0.03	0.09	0.03	0.10	0.03	0.07	0.05	0.05	0.06
	<i>Sk</i>	0.79	0.84	0.72	0.91	1.06	0.80	1.75	0.73	1.22	1.21	1.03	1.28
	<i>Ku</i>	3.45	3.94	3.21	3.67	5.37	3.54	8.46	3.33	5.57	6.18	4.82	6.69
Fin du tableau													

TABLE 3.2 – Comparaison, sur 200 images de la base de BERKELEY, des résultats de la segmentation par la combinaison couleur/texture proposée et des principales méthodes de segmentation de la littérature : (MS = MeanShift, JSEG, FH = Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR = Fusion Clustering Results, CTM = Compression-based Texture Merging, CV = segmentation de CHAN-VESE, CTB segmentation basée sur la combinaison de la couleur, la texture et la luminosité (Color, Texture and Brightness), CF = segmentation basée sur la combinaison fixe, CA = segmentation basée sur la combinaison adaptative, CS = segmentation basée sur la combinaison supremum et C-ACP = segmentation basée sur la combinaison par Analyse en Composantes Principales). Les valeurs en rouge (respectivement en bleue) désignent la première meilleure (respectivement la deuxième meilleure) performance.



FIGURE 3.15 – Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par CTB, les images segmentées par MeanShift, les images segmentées par FCR et les images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER).



FIGURE 3.16 – Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par Ncuts, les images segmentées par CHAN-VESE, les images segmentées par CTM et les images segmentées par Jseg).





FIGURE 3.17 – Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par combinaison adaptative, les images segmentées par combinaison supremum, les images segmentées par combinaison fixe avec le paramètre de pondération  $\alpha$  optimal associé à chaque image et les images segmentées par combinaison par Analyse en Composantes Principales).

Nous constatons également que l'ordre des méthodes de segmentation a quelque peu changé par rapport à l'ordre obtenu dans l'étude comparative effectuée dans le chapitre 2. Par exemple, la segmentation par l'algorithme MeanShift qui était en troisième position est passée en deuxième position. Ceci montre, selon nous, l'intérêt de multiplier les différentes propriétés statistiques pour évaluer la qualité des méthodes de segmentation.

Les figures 3.15, 3.16 et 3.17 illustrent sur cinq images de la base de BERKELEY les résultats de segmentation obtenus avec la méthode de combinaison proposée et les principales méthodes de segmentation de la littérature. Bien que la performance de la combinaison proposée ne soit pas la meilleure, comme démontré dans le tableau 3.2, les résultats obtenus sont très satisfaisants. À titre d'illustration, nous pouvons remarquer pour la troisième image que l'animal est mieux détecté par la méthode proposée que par les méthodes de FELZENSZWALB et HUTTENLOCHER, Ncuts, CHAN-VESE, CTM et les combinaisons adaptative, supremum et fixe. Si nous prenons aussi l'exemple de la première image, les détails et les bordures des deux luges sont mieux identifiés avec la méthode proposée qu'avec le Ncuts, la CTM et la segmentation basée sur une combinaison fixe.

Tout comme nous l'avons fait pour la base de BERKELEY, nous mesurons maintenant la performance de la méthode de segmentation utilisant la combinaison proposée, en fonction des différents couples de gradients couleur et texture utilisés sur la base de texture VISTEX. Cette base ne proposant pas de segmentations de référence, nous présentons dans le tableau 3.3 à l'aide des critères d'évaluation non-supervisée (sans segmentation de référence)  $MSE$ ,  $\mathcal{Z}$  de ZEBOU DJ et  $\mathcal{B}$  de BORSOTTI, la performance de la segmentation, en fonction des différents couples de gradients couleur et texture.

	Combinaison par Analyse en Composantes Principales									
	AM	AZ	GM	GZ	H1M	H1Z	H2M	H2Z	H5M	H5Z
$MSE (\times 10^2)$	4.97	4.13	5.15	4.02	4.94	<b>3.97</b>	<b>5.23</b>	4.48	5.36	3.97
$\mathcal{Z}$	0.41	0.38	0.35	0.38	0.23	0.23	<b>0.42</b>	0.39	0.39	0.38
$\mathcal{B}$	0.19	0.20	0.19	<b>0.18</b>	0.21	0.2	<b>0.18</b>	0.21	<b>0.18</b>	0.19

TABLE 3.3 – Performance de la méthode de segmentation basée sur la combinaison couleur/texture proposée sur la base de texture VISTEX, en fonction des différents couples des gradients couleur et texture.

Au vu des résultats présentés dans le tableau 3.3, nous pouvons constater que le couple de gradients couleur/texture H2M est favorisé par les métriques  $\mathcal{Z}$  et  $\mathcal{B}$ . Nous rappelons que ce même couple s'était déjà distingué lors des tests réalisés sur la base de BERKELEY. Pour cette raison, nous retenons ce couple pour l'étude comparative de la méthode de combinaison couleur/texture proposée avec les principales méthodes de segmentation issues de la littérature et introduites précédemment dans le chapitre 1. Le tableau 3.4 illustre la performance de chaque méthode de segmentation sur les 33 images de la base de texture VISTEX, sélectionnées aléatoirement. Comme pour la base de BERKELEY, nous utilisons les mêmes propriétés statistiques, définies précédemment, pour mesurer la performance de chaque méthode de segmentation. Au vu des résultats fournis dans le tableau 3.4, nous pouvons voir que la segmentation réalisée par l'algorithme MeanShift est la plus performante selon les trois métriques d'évaluation (non-supervisées) utilisées, en considérant la base de texture VISTEX.

		Base de texture VISTEX										
		Méthodes de segmentation										
		MS	JSEG	FH	FCR	NCUTS	CTM	CV	CF	CA	CS	C-ACP
Critères d'évaluation		Évaluation non supervisée										
$MS\mathcal{E}$	<i>Moy</i>	218.7	449.8	422.6	350.8	667.2	671.8	532.3	384.6	389.0	392.8	523.3
	<i>Min</i>	42.3	33.8	96.4	74.9	96.0	107.7	56.5	39.2	53.0	60.6	76.7
	<i>Max</i>	679.2	1436.3	1035.4	848.9	1630.0	2006.7	1572.0	1150.1	1082.7	1059.1	1756.1
	<i>Med</i>	169.1	269.5	346.7	315.4	518.8	544.7	402.4	280.3	316.8	297.7	433.1
	<i>ET</i>	138.1	382.9	256.1	202.9	460.1	494.9	363.4	274.9	257.9	272	392.9
	<i>Sk</i>	1.5	1.2	0.8	0.8	0.8	1.2	1.2	1.2	1.1	1	1.3
	<i>Ku</i>	5.4	3.3	2.8	3.2	2.5	3.6	3.7	3.6	3.4	3	4.3
$\mathcal{Z}$	<i>Moy</i>	0.55	0.15	0.52	0.33	0.27	0.22	0.29	0.42	0.41	0.41	0.42
	<i>Min</i>	0	0	0.22	0	0.05	0.01	0.07	0.04	0.03	0.03	0
	<i>Max</i>	0.81	0.66	0.75	0.75	0.69	0.64	0.49	0.80	0.78	0.71	0.79
	<i>Med</i>	0.56	0.08	0.52	0.34	0.24	0.19	0.29	0.43	0.41	0.44	0.44
	<i>ET</i>	0.18	0.16	0.11	0.18	0.16	0.17	0.11	0.18	0.17	0.17	0.20
	<i>Sk</i>	-0.87	2.01	-0.22	-0.03	1.14	0.85	-0.08	-0.07	-0.23	-0.40	-0.37
	<i>Ku</i>	3.93	6.61	3.21	2.43	3.63	3.01	2.15	2.70	3.01	2.67	2.49
$\mathcal{B}$	<i>Moy</i>	0.06	0.08	0.16	0.08	0.21	0.07	0.24	0.15	0.16	0.15	0.18
	<i>Min</i>	0	0.01	0.01	0	0.02	0.01	0.02	0.01	0.01	0.01	0.02
	<i>Max</i>	0.19	0.18	0.44	0.25	0.55	0.16	0.73	0.45	0.46	0.44	0.52
	<i>Med</i>	0.05	0.09	0.12	0.06	0.16	0.06	0.16	0.12	0.12	0.12	0.14
	<i>ET</i>	0.04	0.05	0.12	0.06	0.15	0.04	0.18	0.11	0.11	0.11	0.13
	<i>Sk</i>	1.10	0.27	1.06	1.18	0.86	0.54	1.19	1.06	1.02	0.98	1.05
	<i>Ku</i>	3.65	2.12	3.04	3.64	2.65	2.53	3.57	3.38	3.25	3.14	3.42

TABLE 3.4 – Comparaison, sur 33 images de la base de VISTEX, des résultats de segmentation basée sur la combinaison couleur/texture proposée et les principales méthodes de segmentation de la littérature : (MS = MeanShift, JSEG, FH = Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR = Fusion Clustering Results, CTM = Compression-based Texture Merging, CV = segmentation de CHAN-VESE, CF = segmentation basée sur la combinaison fixe, CA = segmentation basée sur la combinaison adaptative, CS = segmentation basée sur la combinaison supremum et C-ACP = segmentation basée sur la combinaison par Analyse en Composantes Principales). Les valeurs en rouge (respectivement en bleue) désignent la première meilleure (respectivement la deuxième meilleure) performance.



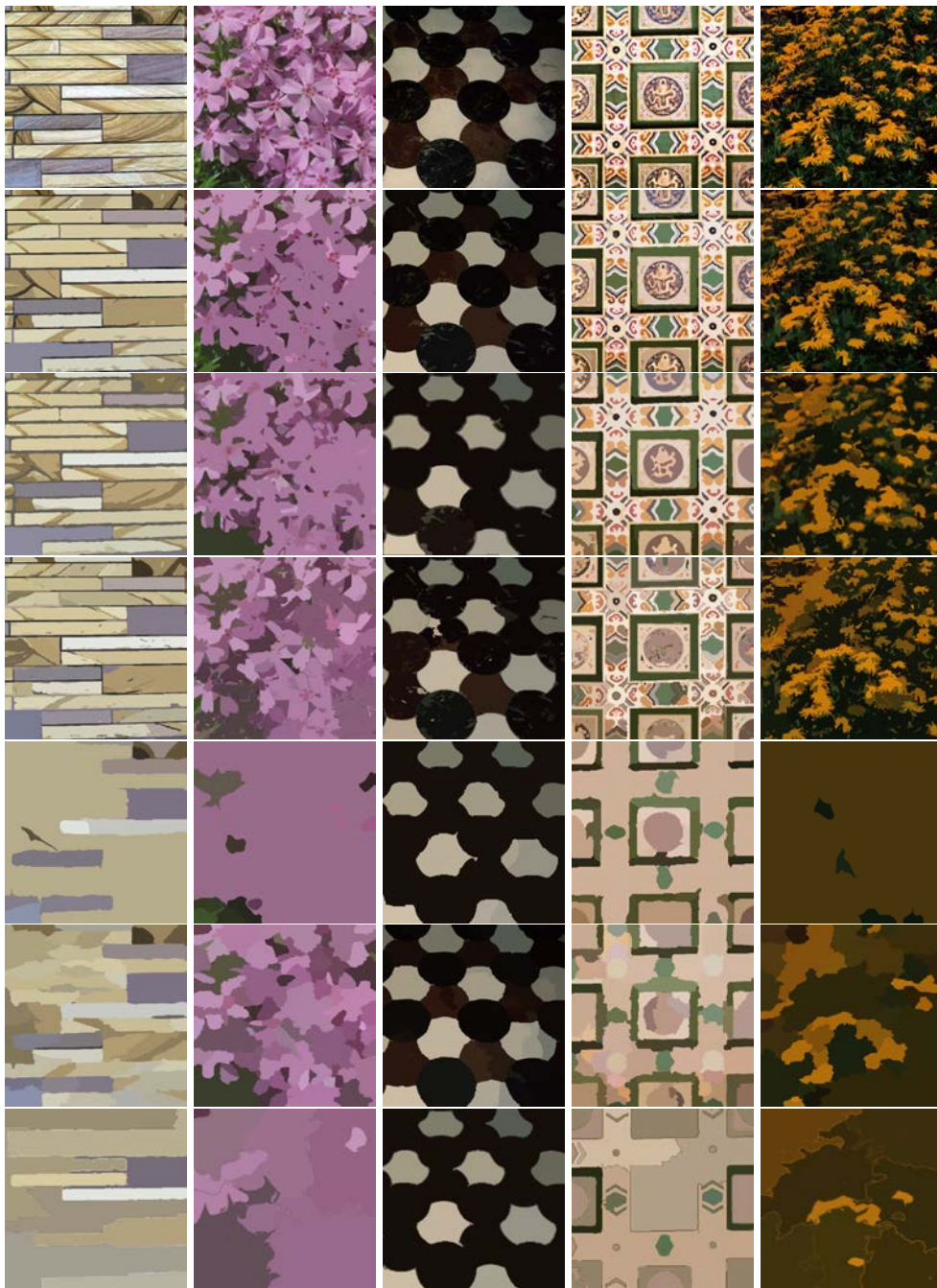


FIGURE 3.18 – Illustration des résultats de segmentation sur quatre images de la base de VISTEX (de haut en bas : les images originales, les images segmentées par MeanShift, les images segmentées par FCR, les images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, les images segmentées par CTM, les images segmentées par Jseg et les images segmentées par Ncuts).

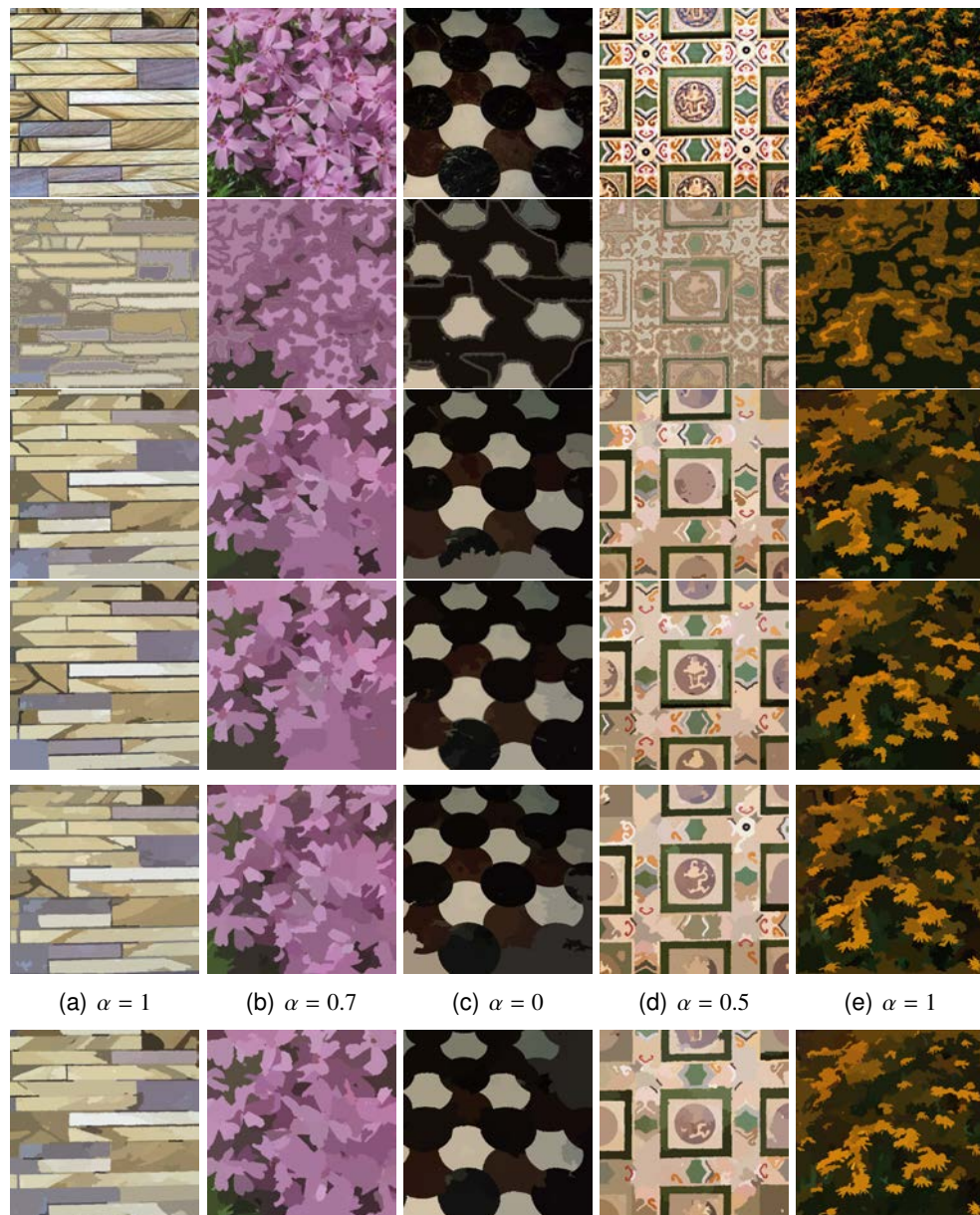


FIGURE 3.19 – Illustration des résultats de segmentation sur quatre images de la base de VISTEX (de haut en bas : les images originales, les images segmentées par la segmentation de CHAN-VESE, les images segmentées par combinaison adaptative, les images segmentées par combinaison supremum, les images segmentées par combinaison fixe avec le paramètre de pondération  $\alpha$  optimal associé à chaque image et les images segmentées par combinaison par comparaison par Analyse en Composantes Principales).

Les figures 3.18 et 3.19 illustrent, sur cinq images sélectionnées de la base de texture VISTEX, les résultats de segmentation basée sur la combinaison proposée et les méthodes de segmentation de la littérature (évaluées dans le tableau 3.4), tout en considérant les meilleurs couples gradients couleur et texture (le couple H5M pour la combinaison fixe avec le paramètre de pondération  $\alpha$  optimal associé à chaque image et les images segmentées par combinaison par comparaison par Analyse en Composantes Principales).



qu'il est assez difficile d'identifier les meilleurs algorithmes de segmentation (hormis pour le MeanShift), étant donné que les métriques utilisées ne les placent pas dans la même position. En effet, prenons l'exemple de CTM qui arrive en deuxième position avec la métrique de BORSOTTI  $\mathcal{B}$  alors qu'il se positionne qu'en dernière position avec la métrique  $MSE$ . Il en est de même pour la méthode FCR qui se positionne en deuxième position avec  $MSE$  et en septième position avec la métrique de ZEBOUJ  $\mathcal{Z}$ . De même, la méthode de FELZENSZWALB et HUTTENLOCHER arrive en deuxième position avec la métrique de ZEBOUJ  $\mathcal{Z}$  mais en huitième position avec le critère de BORSOTTI  $\mathcal{B}$ . Notons que la méthode de segmentation par la combinaison couleur/texte basée sur l'ACP oscille entre la quatrième et la neuvième position, selon la métrique d'évaluation choisie, ce qui reste toutefois acceptable. Ce résultat peut s'expliquer par le type d'images présentes dans la base de texture VISTEX. En effet, cette base est dédiée aux images texturées et ne présente pas de scènes complexes avec une multitude de couleurs. Notre méthode de combinaison couleur/texte ne peut donc pas exercer pleinement son potentiel, comme elle a pu le faire pour la base de BERKELEY.

### 3.6/ CONCLUSION

Dans ce chapitre, nous avons proposé et développé une nouvelle méthode de combinaison des informations de couleur et de texture pour la segmentation d'images couleur. La méthode proposée est basée sur l'étude des axes principaux d'un ensemble de points 3D formé par des descripteurs couleur et texture, définis dans le chapitre 1. La nouveauté de cette technique réside dans la définition d'un gradient structurel à l'aide d'une combinaison hybride et non paramétrique à partir d'un ensemble de gradients couleur et texture. La technique de combinaison proposée s'adapte localement au contenu de l'image puisqu'elle permet d'attribuer à chaque pixel l'information la plus pertinente. En plus, l'approche proposée ne nécessite aucun paramétrage, contrairement à d'autres méthodes de la littérature. La segmentation par l'algorithme LPE utilisant le gradient structurel, issu de la combinaison couleur/texte proposée, et couplé à une réduction adéquate de germes, permet d'obtenir dans l'espace  $(R, G, B)$  de bons résultats de segmentation. En effet, les résultats de segmentation basée sur cette nouvelle combinaison couleur/texte ont été évalués et comparés aux résultats des principales méthodes de segmentation de la littérature connues par leurs performances. Nous avons montré qu'au delà de l'avantage de ne pas nécessiter de paramétrage, la méthode de combinaison proposée, bien qu'offrant des résultats moyens sur la base d'images de texture VISTEX (sa position oscille entre la 4<sup>ème</sup>/11 et 9<sup>ème</sup>/11 selon la métrique utilisée), offre des résultats très satisfaisants sur la base d'images couleur de BERKELEY (la méthode se situe en moyenne en 6<sup>ème</sup> position sur 12 méthodes testées). Nous avons montré également que lorsque la méthode proposée n'occupe pas la première ou la deuxième position (suivant la métrique choisie), l'écart numérique des valeurs des propriétés statistiques par rapport aux méthodes de segmentation qui la précèdent est relativement faible. Enfin, nous avons également montré dans ce chapitre, l'importance de la diversité des descripteurs couleur et texture. Ainsi, nous avons pu constater que le meilleur gradient structurel est obtenu par combinaison d'un gradient morphologique couleur et d'un gradient de texture basé sur le deuxième paramètre d'HARALICK  $H_2$  (le contraste). Bien évidemment, une perspective à court terme consiste à augmenter le nombre de descripteurs et à s'intéresser davantage au principe de fusion des informations couleur/texte.

# CARACTÉRISATION DE L'ENVIRONNEMENT DE RÉCEPTION DES SIGNAUX GNSS

## Sommaire

<b>4.1</b>	<b>Introduction</b>	<b>141</b>
<b>4.2</b>	<b>État de l'art</b>	<b>142</b>
<b>4.3</b>	<b>Les projets ViLoc, CAPLOC</b>	<b>144</b>
<b>4.4</b>	<b>Méthode d'évaluation de la caractérisation de l'environnement</b>	<b>147</b>
<b>4.5</b>	<b>Stratégie de détection du ciel</b>	<b>148</b>
4.5.1	Première approche : Segmentation par combinaison couleur/texture et classification de régions	148
4.5.2	Deuxième approche : Simplification de l'image et classification pixellaire	154
4.5.3	Classification pixellaire	165
<b>4.6</b>	<b>Paramétrage de la seconde approche proposée</b>	<b>170</b>
<b>4.7</b>	<b>Conclusion</b>	<b>180</b>

## 4.1/ INTRODUCTION

Dans les environnements contraints, notamment urbains, les signaux GNSS (Global Navigation Satellites Systems) peuvent ne pas être reçus directement et donc subir des réflexions, voire être totalement bloqués par les obstacles constituant l'environnement de réception (infrastructures, végétation, etc). Ce phénomène, illustré sur la figure 4.1, peut fausser l'information de positionnement suite aux retards engendrés par les multi-trajets. Nous verrons dans le chapitre suivant que l'utilisation conjointe de l'image et des données GNSS peut être une solution pour détecter les satellites dont les signaux sont reçus directement (respectivement après réflexions) et pondérer ainsi certains d'entre-eux dans le calcul d'une nouvelle position, se voulant plus précise. Avant d'aborder cette partie, il convient de caractériser l'environnement de réception des signaux satellitaires en effectuant une analyse d'images acquises en mobilité à l'aide d'une caméra installée sur le toit d'un véhicule (en mouvement) et orientée vers le ciel. À ce stade, il s'agit d'identifier les régions " ciel " (respectivement les régions " non ciel ") de l'image, dans l'objectif de détecter les satellites situés dans ces différentes régions après la projection de leur position (étape décrite dans le chapitre suivant). Pour cette raison,

deux approches sont alors proposées, dans ce chapitre. La première approche consiste à effectuer une segmentation de l'image par combinaison des informations de couleur et de texture, présentée dans le chapitre 3, puis à classer les régions obtenues en deux classes via un algorithme de classification adapté. La deuxième approche proposée est basée sur une simplification de l'image par un filtre adapté, suivie d'une classification pixellaire en deux classes afin d'obtenir une zone " ciel " et une zone " non ciel ". Les tests expérimentaux présentés dans ce chapitre sont effectués sur deux bases d'images réelles acquises en mobilité via des plateformes de laboratoire, dans le cadre de deux projets, LOCOPROL et CAPLOC, introduits dans la suite de ce chapitre.

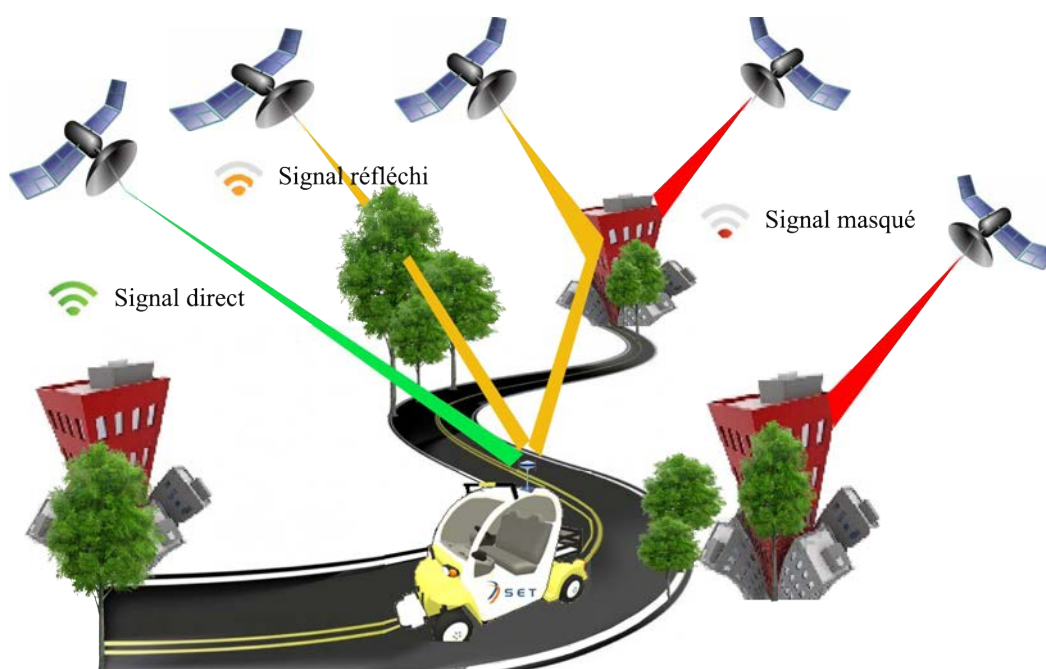


FIGURE 4.1 – Illustration du phénomène de multi-trajets en milieu urbain.

Ce chapitre est organisé de la façon suivante. Dans un premier temps, différents travaux de la littérature utilisant la vision par ordinateur pour des applications liés aux systèmes de localisation seront présentés brièvement. Dans un deuxième temps, la stratégie globale ainsi que l'intérêt de la caractérisation de l'environnement de réception des signaux satellitaires seront positionnés dans le cadre du projet CAPLOC. Nous développons ensuite les deux approches proposées et présentons les résultats expérimentaux sur deux bases de données réelles, pour conclure sur la stratégie de caractérisation de l'environnement de réception des signaux GNSS qui sera utilisée dans le chapitre 5 pour améliorer le calcul de la position d'un mobile dans un environnement urbain contraint.

## 4.2/ ÉTAT DE L'ART

Durant ces dernières années, nous avons observé une forte croissance des travaux utilisant la vision par ordinateur pour des applications liées aux systèmes de localisation. Ces travaux peuvent être différenciés en fonction du type de caméras utilisées ou en fonction de la façon dont l'image est exploitée dans les systèmes de localisation. Sur le

premier point, nous pouvons distinguer les méthodes utilisant des images perspectives [Robertson 2004, Zhang 2006, Royer 2007, Jacobs 2007, Hays 2008] et d'autres méthodes utilisant des images omni-directionnelles permettant ainsi d'avoir un champ de vision plus large [Stein 1995, Koch 2007, Meguro 2009, Tardif 2008, Ramalingam 2010]. Si l'on considère la façon dont l'image est exploitée dans les systèmes de localisation, nous pouvons citer des méthodes basées sur des algorithmes d'appariement qui consistent à trouver, à partir d'une base de données d'images géoréférencées (par un GPS), l'image la plus proche de l'image requête [Robertson 2004, Zhang 2006, Jacobs 2007, Hays 2008]. Une autre solution consiste à utiliser des images (principalement omni-directionnelles) souvent couplées à des modèles 3D pour détecter la ligne d'horizon séparant le ciel des autres objets constituant l'environnement [Meguro 2009, Ramalingam 2010].

Dans cette thèse, nous nous sommes focalisés sur la deuxième catégorie de méthodes, qui utilisent l'image pour détecter la ligne d'horizon. La détection de cette dernière revient à déterminer la frontière entre les zones " ciel " et " non ciel " de l'image. Nous décrirons brièvement le principe de quelques méthodes existantes de cette catégorie, afin de mieux situer l'intérêt de l'approche proposée dans ce chapitre.

KOCH et TELLER [Koch 2007] proposent une méthode de localisation qui utilise un modèle 3D connu de la scène étudiée et une caméra grand angle. Des scènes d'intérieur ont été utilisées dans le cadre de cette méthode. La technique consiste à appairer les lignes existantes dans le modèle 3D avec les lignes correspondantes dans la séquence d'images omni-directionnelles. La mise en correspondance est effectuée entre les contours présents dans l'image du monde réel et les segments issus du modèle 3D. L'hypothèse de ce système est de considérer que l'environnement présente un nombre important de lignes droites. La solution proposée est constituée de deux phases : la première concerne la phase d'initialisation de la pose de la caméra et requiert une minute pour un volume de recherche de trois mètres. Cela s'effectue par une première mise en correspondance entre les pixels-contours dans l'image et les segments dans le modèle 3D. Quant à la deuxième phase, elle relève d'une phase de maintenance. Cette solution a été testée sur des séquences d'images de taille  $512 \times 384$  pixels ; le temps de calcul atteignant 10 secondes/image. Cette solution nous semble inadaptée au vu de la contrainte temps-réel requise pour notre application. MEGURO et al. [Meguro 2009] proposent une méthode qui vise l'amélioration de la précision du positionnement. Les auteurs utilisent une caméra infrarouge en visant l'élimination des satellites NLOS (Non-line-Of-Sight), considérés comme un facteur dégradant de la performance du système de positionnement. En effet, grâce à la caméra infrarouge, le ciel apparaît particulièrement sombre, ce qui facilite la détection de la frontière entre le ciel (pixels sombres) et les bâtiments (pixels blancs). Les inconvénients majeurs de cette approche résident tout d'abord dans le coût élevé du système rendant impossible son déploiement à grande échelle. De plus, l'utilisation de ce type particulier de caméra demande une méthode de projection précise de la position des satellites. Un étalonnage spécifique avec des sources ponctuelles thermiques disposées à l'intérieur est requis, ce qui rend complexe une fois encore l'exploitation de cette solution. Enfin, selon les auteurs, la détection de la ligne d'horizon est principalement basée sur une projection azimutale, mais le résultat obtenu ne reflète pas la ligne d'horizon réelle et l'approche peut potentiellement mener à des erreurs lorsque la caméra n'est pas orientée verticalement. RAMALINGAM



et al. [Ramalingam 2010] proposent quant à eux une méthode qui vise la détection de la ligne d'horizon en utilisant une caméra omni-directionnelle dans le domaine du visible, au lieu d'une caméra infrarouge, ainsi qu'un modèle 3D de l'environnement. La technique proposée consiste, dans un premier temps, à calibrer la caméra à objectif fisheye afin d'estimer ses paramètres intrinsèques. Une détection de contours est ensuite réalisée afin d'extraire une première ligne d'horizon dite synthétique. Ceci est réalisé en utilisant l'image fisheye synthétique, obtenue à l'aide des paramètres de calibrage trouvés précédemment, et le modèle 3D. D'autre part, une classification de l'image réelle, acquise à l'aide de la caméra fisheye, est effectuée via la résolution d'un problème de minimisation d'énergie afin de détecter la deuxième ligne d'horizon dite réelle. La troisième et dernière phase consiste à mettre en correspondance les deux résultats de ligne d'horizon (synthétique et réelle). Nous citons finalement COHEN et al. [Cohen 2009, Cohen 2010b] qui caractérisent l'environnement de réception de signaux GNSS en s'appuyant sur une segmentation adaptative des informations de couleur et de texture, couplée à une classification pixellaire binaire (en deux classes). C'est cette orientation (combinaison couleur/texture) que nous avons suivie dans le cadre du développement de notre première approche de caractérisation de l'environnement de réception des signaux GNSS.

### 4.3/ LES PROJETS ViLoc, CAPLOC

Avant de détailler la stratégie proposée pour caractériser l'environnement de réception des signaux GNSS, il nous semble important de donner une brève présentation du contexte des travaux développés dans le présent chapitre et le chapitre suivant : les projets ViLoc et CAPLOC.

Dans le domaine du transport de voyageurs et de fret, la volonté des opérateurs est de réduire les équipements attachés à l'infrastructure notamment en raison de l'augmentation des actes de vandalisme, que subissent ces matériels, et de la hausse des coûts de leur maintenance. Aujourd'hui, les systèmes GNSS (Global Navigation Satellite System) constituent une réponse très adaptée à cette problématique. Parmi ces systèmes et en attente de la solution européenne Galileo, le système de radionavigation par satellites Américain GPS (Global Positioning System) offre le meilleur compromis coût/simplicité/performance ; par ailleurs, les performances d'un récepteur GPS sont améliorées lorsque les services EGNOS (European Geostationary Navigation Overlay Service) sont utilisés.

De nombreuses études ont cependant montré que le service de localisation est sensiblement dégradé en environnement urbain. Dans ces milieux, les conditions de propagation des signaux sont fortement altérées par la présence et la densité des obstacles situés autour de l'antenne de réception. Les signaux subissent des réflexions multiples qui affectent sensiblement les résultats fournis par l'algorithme d'estimation de position. Les services EGNOS, portés par la constellation de satellites géostationnaires dont la réception est très difficile dans les canyons urbains, sont également souvent inexploitable pour améliorer la précision de l'information de localisation.

L'objectif du projet CAPLOC (projet PREDIT GO2 2010-2014 faisant intervenir l'IFSTTAR-LEOST et l'UTBM/IRTES-SET) est de développer un système capable d'améliorer la précision de localisation d'un récepteur GNSS dans les milieux urbains à partir de la connaissance de la structure 3D de l'environnement traversé par le mobile. L'ambition n'est pas de développer un nouveau capteur mais d'étudier comment il est possible d'améliorer les performances des récepteurs grand publics actuels à partir de l'analyse des images fournies par deux caméras grand angle installées sur le toit du mobile. Pour y parvenir, deux axes sont étudiés. Dans un premier temps, un algorithme de détection temps-réel du ciel et des bâtiments permet de déterminer les zones de l'espace à partir desquelles un signal ne peut pas être reçu par trajet direct. Cette fonction s'inspire de ce que propose un récepteur RAIM (Receiver Autonomous Integrity Monitoring) utilisé dans l'aviation et qui permet d'exclure les satellites dégradant les performances du récepteur. C'est cette fonction limitée à la détection temps-réel des zones " ciel " et " non ciel " de l'image qui sera détaillée dans ce chapitre et utilisée pour le calcul d'une nouvelle position dans le chapitre 5 de ce manuscrit. Dans un second temps, il s'agit de calculer un modèle 3D suffisamment fidèle de l'environnement qui, intégré à un logiciel de lancer de rayons, permettra de corriger les erreurs de pseudodistance causées par les réflexions multiples. Le projet ViLoc (projet REGION 2009-2013 faisant intervenir l'IFSTTAR-LEOST et l'UTBM/IRTES-SET) a été proposé pour initier une collaboration entre les deux établissements sur le sujet de localisation précise d'un mobile et ainsi prendre de l'avance de phase sur le projet CAPLOC. L'objectif du projet ViLoc a été de caractériser, par l'image, l'environnement de réception des signaux GNSS sans tenir compte de la contrainte temps-réel. Ceci a permis de détecter la ligne d'horizon et d'identifier les satellites dont les signaux sont reçus via des trajets directs et via des trajets réfléchis. Ces données sont ensuite réutilisées et améliorées dans le cadre du projet CAPLOC afin de tenir compte de la contrainte temps-réel (présent chapitre) et de proposer des méthodes de calcul de position plus précise par exclusion ou pondération de l'usage des satellites (chapitre 5).



FIGURE 4.2 – Véhicules instrumentés de l'UTBM/IRTES-SET (à gauche) et de l'IFSTTAR-LEOST (à droite).

Dans le projet CAPLOC, nous utilisons un système complet d'acquisition d'images, installé sur un véhicule de laboratoire, permettant de réaliser des acquisitions de données

en mobilité, dans des conditions réelles d'utilisation. Ceci permet de se confronter aux difficultés liées aux applications embarquées (contrainte temps-réel, surexposition de l'image, changements brusques de luminosité, etc.). Dans le cadre de nos travaux, nous utilisons une caméra monoCCD couleur, avec un objectif fisheye, installée sur le toit du véhicule de laboratoire et orientée vers le ciel. Plusieurs récepteurs GNSS sont également utilisés : un récepteur bas-coût et un récepteur Septentrio afin de récupérer l'information de position et les données brutes, ainsi qu'un récepteur GNSS-RTK pour disposer d'une trajectoire de référence. Un Pc-Rack a été également embarqué à bord du véhicule permettant la sauvegarde des images lors de l'acquisition. La figure 4.2 illustre le système d'acquisition installé sur le véhicule de l'UTBM/IRTES-SET et sur le camion de laboratoire de l'IFSTTAR-LEOST.

Dans ce chapitre, nous testerons et évaluons les deux approches de caractérisation de l'environnement de réception des signaux GNSS proposées sur deux bases de données images. La première, illustrée par cinq images sur la figure 4.3, a été acquise dans le cadre du projet LOCOPROL (en décembre 2004) et nous permettra, suite à une série de tests expérimentaux, d'obtenir des premières conclusions sur les deux meilleurs algorithmes de simplification d'images.

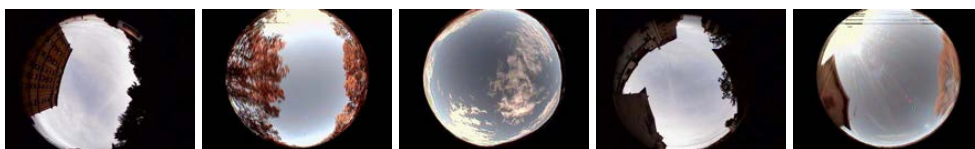


FIGURE 4.3 – Illustration de cinq images issues de la base du projet LOCOPROL.

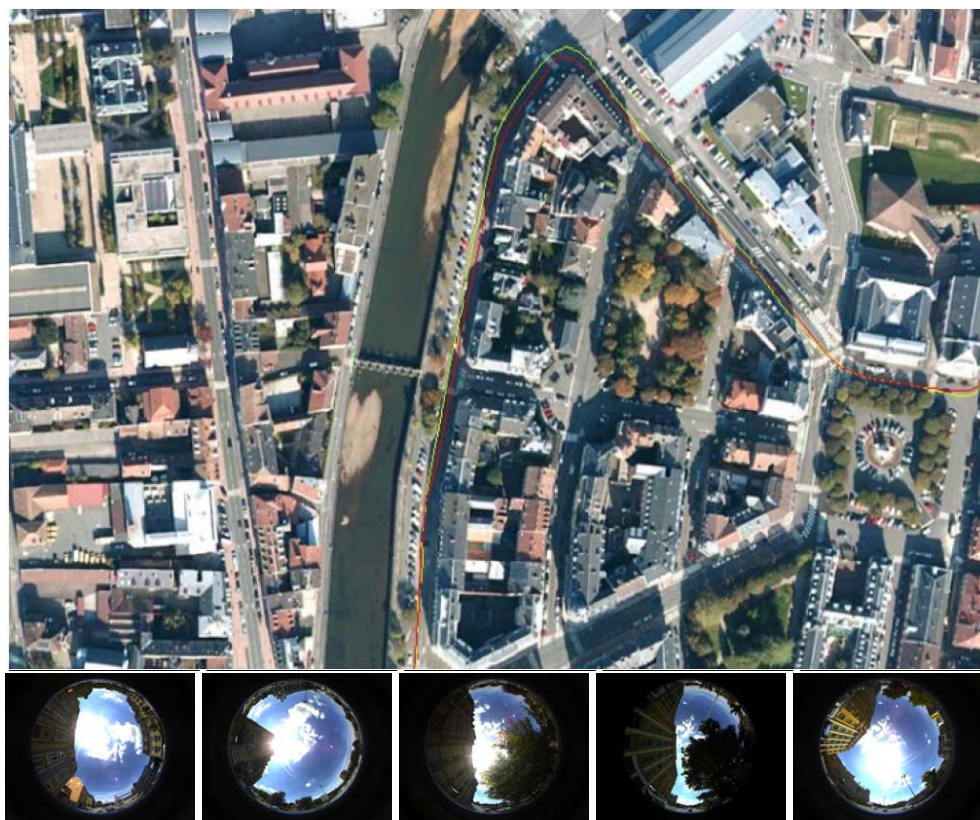


FIGURE 4.4 – Illustration d'un trajet d'une campagne d'acquisition réalisée à Belfort, dans le cadre du projet CAPLOC avec 5 des 149 images acquises en mobilité.



La seconde base acquise dans le cadre du projet CAPLOC, nous permettra d'optimiser les paramètres des algorithmes implémentés et de prouver la robustesse de la stratégie proposée. Cette base, réalisée en juin 2010 à Belfort, dispose de données hétérogènes et de complexités variées (canyon urbain, abondance de végétation, surexposition, changements brusques de luminosité, etc.). Elle constitue, de ce fait, pour la validation de notre stratégie de caractérisation de l'environnement de réception GNSS. La figure 4.4 illustre un des trajets d'une campagne d'acquisition effectuée à Belfort avec cinq des 149 images de cette base, qui a été utilisée pour optimiser puis évaluer les deux approches proposées.

#### 4.4/ MÉTHODE D'ÉVALUATION DE LA CARACTÉRISATION DE L'ENVIRONNEMENT

Les résultats de caractérisation de l'environnement de réception des signaux GNSS obtenus via les deux approches proposées dans ce chapitre ont été évalués à l'aide de la méthode définie ci-dessous. Cette méthode fait appel à une classification de référence réalisée par un expert et fait intervenir cinq mesures que nous décrivons ci-après. Soit  $I_o^c$  (respectivement  $R_o$ ) la zone d'une image classée  $I^c$  (respectivement d'une classification de référence) délimitée par un masque qui correspond au champ de vision de la caméra. Soit  $I_o$  la zone de l'image initiale  $I$  délimitée par le contour du masque de l'objectif de la caméra. Les cinq métriques d'évaluation sont définies comme suit.

- 1) Le pourcentage de pixels faisant partie de la zone " ciel " dans les deux images (classée et de référence) est donné par l'équation suivante :

$$C_{I^c}C_R = \frac{\text{card}\{p/ p \in I_o^c, p \in R_o\}}{\text{card}\{p/ p \in I_o\}} \times 100 \quad (4.1)$$

- 2) Le pourcentage de pixels classés comme " ciel " dans l'image classée et ne faisant pas partie de la zone " ciel " dans l'image de référence est donné par l'équation suivante :

$$C_{I^c}\bar{C}_R = \frac{\text{card}\{p/ p \in I_o^c, p \notin R_o\}}{\text{card}\{p/ p \in I_o\}} \times 100 \quad (4.2)$$

- 3) Le pourcentage de pixels classés comme " non-ciel " dans les deux images (classée et de référence) est donné par l'équation suivante :

$$\bar{C}_{I^c}\bar{C}_R = \frac{\text{card}\{p/ p \notin I_o^c, p \notin R_o\}}{\text{card}\{p/ p \in I_o\}} \times 100 \quad (4.3)$$

- 4) Le pourcentage de pixels classés comme " non-ciel " dans l'image classée et qui ne font pas partie de la région " ciel " dans l'image de référence est donné par l'équation suivante :

$$\bar{C}_{I^c}C_R = \frac{\text{card}\{p/ p \notin I_o^c, p \in R_o\}}{\text{card}\{p/ p \in I_o\}} \times 100 \quad (4.4)$$

- 5) Une dernière mesure traduit le pourcentage de pixels correctement classés (appelée "Taux de Bonne Classification : TBC" dans la suite de ce rapport) et est donnée par l'équation suivante :

$$TBC = C_{I^c}C_R + \bar{C}_{I^c}\bar{C}_R \quad (4.5)$$

## 4.5/ STRATÉGIE DE DÉTECTION DU CIEL

Dans cette section, nous commençons par présenter le synopsis de notre stratégie de détection des zones de "ciel" et de "non-ciel" dans une image fisheye. Le résultat de cette détection permettra, après repositionnement des satellites dans l'image, d'exclure ou pondérer l'impact de chacun d'entre-eux dans le calcul de position de localisation (cf. chapitre 5). Le synopsis illustré sur la figure 4.5 est composé tout d'abord d'une brique de segmentation ou de simplification de l'image, selon l'approche utilisée. Une deuxième brique concerne l'étape de classification de pixels/régions de l'image à l'aide d'un classifieur adapté. Une brique de post-traitement permet enfin de reclasser les pixels erronés puis de disposer d'une classification binaire de l'image en deux classes "ciel" et "non ciel".

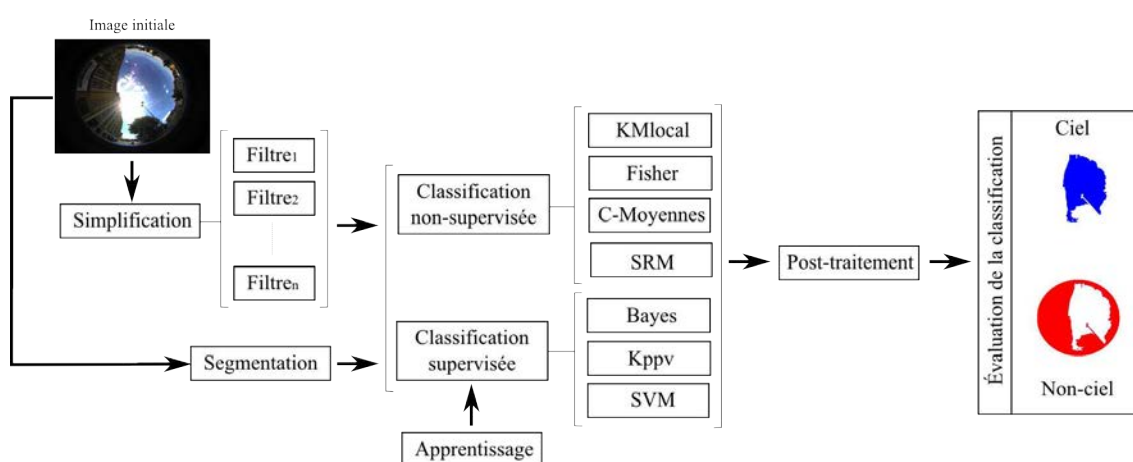


FIGURE 4.5 – Récapitulatif des étapes de la stratégie de caractérisation de l'environnement de réception des signaux GNSS.

Dans les sections suivantes, nous présentons les deux approches de détection de zones "ciel"/"non ciel" développées dans le cadre de cette thèse. La première approche, décrite ci-après, est basée sur une segmentation de l'image par combinaison des informations de couleur et de texture, couplée à une classification de régions. La seconde approche, décrite dans la section d'après, est basée sur une simplification d'image, suivie d'une classification faite cette fois au niveau du pixel. Une étape de post-traitement est ajoutée à la deuxième approche afin de corriger la classification de pixels erronés. Cette deuxième approche a l'avantage, comme on va le voir, de respecter la contrainte temps-réel imposée par l'application.

### 4.5.1/ PREMIÈRE APPROCHE : SEGMENTATION PAR COMBINAISON COULEUR/TEXTURE ET CLASSIFICATION DE RÉGIONS

La première approche de caractérisation de l'environnement de réception des signaux GNSS consiste initialement à appliquer la méthode de segmentation basée sur la combinaison couleur/texture par comparaison de valeurs propres. Plus précisément, il s'agit de procéder à une croissance de régions à l'aide de l'algorithme de Ligne de Partage des Eaux (LPE), en utilisant comme image potentiel le gradient hybride obtenu

suite à la combinaison couleur/texture. Cette combinaison, développée et détaillée dans le chapitre 3, permet d'exploiter de façon optimale les caractéristiques texturales et radiométriques de l'image. Ce traitement nous permet d'obtenir les différentes régions d'intérêt de l'environnement de réception des signaux GNSS que nous classons ensuite en deux classes (" ciel "/" non ciel "). Les tests présentés ci-après ont été effectués sur deux bases d'images issues des projets LOCOPROL et CAPLOC.

Le tableau 4.1 présente le taux de bonne classification moyenne (TBC précédemment défini) de " ciel " et de " non ciel ", en utilisant les images de la base LOCOPROL. La classification binaire est effectuée avec l'algorithme de Fisher. Le taux de bonne classification est présenté, dans le tableau 4.2, en fonction des couples gradients couleur/texture combinés selon trois méthodes de combinaison à savoir la combinaison fixe, la combinaison hybride (adaptative et supremum) et la combinaison basée sur l'Analyse en Composantes Principales (proposée dans le chapitre 3). Nous rappelons au lecteur que la désignation des notations utilisées pour les couples gradients couleur/texture sont présentées dans le tableau 2.2 du chapitre 2.

D'après le tableau 4.1, nous pouvons constater que le meilleur taux de bonne classification de " ciel " / " non ciel " est, dans la majorité des cas, obtenu par la combinaison couleur/texture basée sur l'Analyse en Composantes Principales. En effet, celle-ci se positionne 21 fois sur 40 en première position (contre 12 fois sur 40 pour la combinaison fixe et 7 fois sur 40 pour la combinaison hybride) et ne requiert pas de paramètre de pondération, contrairement à la combinaison fixe [Angulo 2007]. Nous constatons également que le meilleur taux de bonne classification, atteignant 95.7%, est obtenu à chaque fois sur la composante bleue de l'espace couleur ( $R, G, B$ ).

La figure 4.6 illustre les résultats de segmentation (obtenus par combinaison couleur/texture fixe, hybride et proposée) et de classification (obtenus par l'algorithme Fisher) sur cinq images issues de la base LOCOPROL. Pour une meilleure visualisation, seuls les résultats avec les paramètres optimaux (le paramètre de pondération et le gradient couleur/texture) sont présentés. Nous pouvons constater que les résultats obtenus avec la segmentation basée sur la combinaison couleur/texture proposée, même s'ils restent très proches, sont légèrement meilleurs sur la première, la troisième et la quatrième images de la figure 4.6. En effet, les contours des objets d'intérêt sont mieux détectés. Notons aussi qu'aucune des trois méthodes n'est capable de segmenter puis classer la végétation comme " non ciel ", sur la cinquième image.

Le tableau 4.2 présente le taux de bonne classification moyen du " ciel " et du " non ciel ", en utilisant sur les 149 images de la base CAPLOC. Tout comme sur la base LOCOPROL, la classification binaire est effectuée avec l'algorithme de classification Fisher. Le taux de bonne classification est présenté dans le tableau 4.2, en fonction des couples gradients couleur/texture combinés via trois méthodes de combinaison à savoir la combinaison fixe, la combinaison hybride (adaptative et supremum) et la combinaison proposée basée sur l'ACP.



		Base d'images LOCOPROL							
		Combinaison fixe					Combinaison hybride		Combinaison proposée
		0	0.2	0.5	0.7	1	adaptative	supremum	
AM	R	82.30	88.27	89.04	89.40	89.67	89.40	90.26	<b>92.19</b>
	G	84.00	92.31	94.07	94.90	93.71	<b>94.94</b>	94.53	94.66
	B	84.12	94.28	94.61	<b>95.41</b>	95.03	94.95	94.76	94.97
	Moy	93.69	89.99	94.08	92.12	93.69	93.29	92.78	<b>94.61</b>
AZ	R	82.30	88.83	89.92	89.75	89.57	89.12	89.65	<b>91.04</b>
	G	84.00	93.62	94.39	94.13	94.50	<b>94.83</b>	94.63	94.15
	B	84.12	93.84	94.65	94.85	94.65	94.82	<b>94.88</b>	94.34
	Moy	90.78	92.66	91.59	<b>93.96</b>	90.78	93.63	93.72	93.74
GM	R	85.93	87.67	89.93	89.38	89.67	90.17	90.29	<b>92.68</b>
	G	89.20	92.09	92.66	94.28	93.71	94.46	94.78	<b>95.19</b>
	B	89.69	92.06	94.61	94.65	95.03	94.43	94.69	<b>95.35</b>
	Moy	93.69	89.39	92.69	94.12	93.69	93.58	94.05	<b>95.19</b>
GZ	R	85.93	87.44	88.81	89.63	89.57	90.01	89.75	<b>91.59</b>
	G	89.20	90.99	93.96	94.47	94.50	94.98	<b>95.22</b>	94.68
	B	89.69	93.12	94.29	95.42	94.65	<b>95.51</b>	95.37	94.88
	Moy	90.65	90.44	93.85	93.71	90.65	94.27	<b>94.73</b>	94.26
H1M	R	84.83	77.36	87.07	<b>90.05</b>	89.67	89.62	87.25	88.42
	G	90.50	83.91	89.93	<b>95.26</b>	93.71	92.16	90.98	93.42
	B	91.22	84.41	90.23	<b>95.35</b>	95.03	92.45	91.38	93.71
	Moy	93.69	80.27	89.92	<b>95.20</b>	93.69	91.88	90.92	92.57
H1Z	R	84.83	82.06	87.88	88.64	89.57	88.80	89.38	<b>89.79</b>
	G	90.50	86.72	91.13	<b>95.17</b>	94.50	92.89	92.04	93.93
	B	91.22	86.65	91.24	<b>95.39</b>	94.65	92.04	92.10	94.03
	Moy	90.65	86.26	90.79	93.64	90.65	92.48	91.55	<b>93.81</b>
H2M	R	86.29	87.18	89.96	89.68	89.67	90.08	90.99	<b>92.92</b>
	G	87.15	91.51	94.52	95.00	93.71	94.91	94.18	<b>95.33</b>
	B	87.52	91.67	95.04	95.06	95.03	95.36	95.18	<b>95.44</b>
	Moy	93.69	89.71	93.60	93.53	93.69	93.80	93.50	<b>95.00</b>
H2Z	R	86.29	89.73	90.23	90.21	89.57	90.07	89.70	<b>90.66</b>
	G	87.15	94.02	<b>95.10</b>	<b>95.10</b>	94.50	94.69	95.01	94.64
	B	87.52	94.46	95.62	95.46	94.65	95.30	<b>95.67</b>	94.89
	Moy	90.65	93.62	92.09	93.31	90.65	93.26	93.18	<b>93.75</b>
H5M	R	90.72	90.43	90.08	89.89	89.67	89.19	89.34	<b>92.34</b>
	G	94.11	94.45	94.19	95.02	93.71	94.09	94.05	<b>95.25</b>
	B	94.55	94.46	94.17	95.10	95.03	95.06	94.06	<b>95.48</b>
	Moy	93.69	91.72	93.88	93.14	93.69	92.25	92.85	<b>94.97</b>
H5Z	R	90.72	89.69	89.72	89.97	89.57	90.23	90.44	<b>92.26</b>
	G	94.11	93.72	94.23	<b>94.95</b>	94.50	93.96	94.39	94.65
	B	94.55	94.95	94.83	<b>95.27</b>	94.65	95.11	94.95	94.89
	Moy	90.65	93.28	<b>93.90</b>	91.53	90.65	93.56	92.50	93.55

TABLE 4.1 – Résultats de classification binaire (" ciel "/" non ciel ") en fonction de la méthode de combinaison couleur/texture utilisée (fixe, hybride et proposée) et des couples gradients couleur/texture, sur la base d'images du projet LOCOPROL.

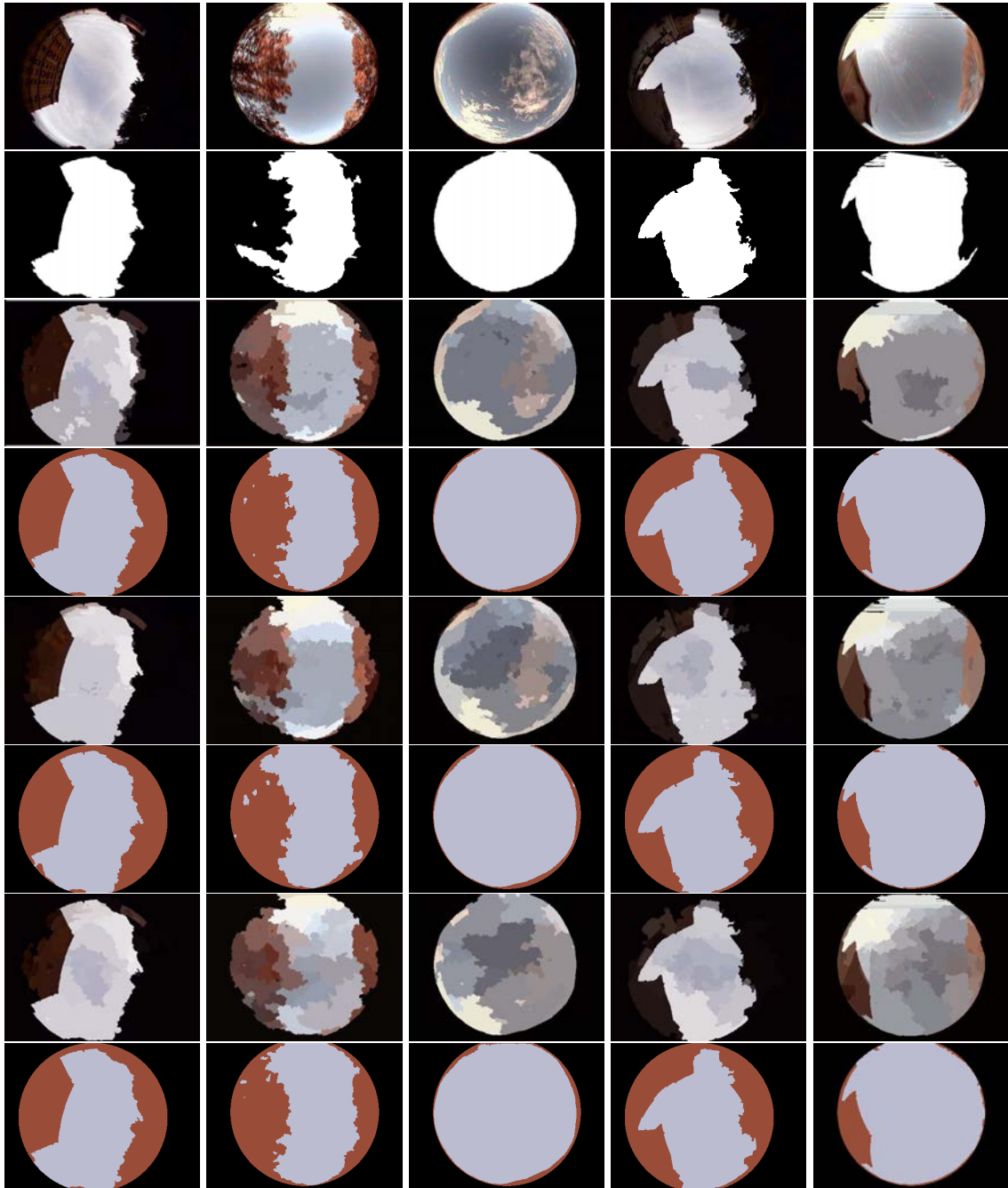


FIGURE 4.6 – Résultats de segmentation et de classification obtenue avec l'algorithme Fisher pour chaque combinaison couleur/texture sur une sélection d'images de la base du projet LOCOPROL (de haut en bas : images originales, vérités-terrain, images segmentées puis classées par la combinaison fixe, la combinaison hybride et par la combinaison proposée).

		Base d'images CAPLOC							
		Combinaison fixe					Combinaison hybride		Combinaison proposée
		0	0.2	0.5	0.7	1	adaptative	supremum	
AM	R	83.88	87.41	86.91	86.39	87.67	88.43	87.27	<b>90.59</b>
	G	88.52	94.76	95.59	95.68	<b>95.79</b>	95.69	95.77	95.64
	B	90.39	96.13	97.68	97.74	97.63	97.67	<b>97.8</b>	96.83
	Moy	89.06	94.94	95.78	95.78	95.81	95.75	<b>95.84</b>	95.73
AZ	R	83.83	85.59	86.01	85.76	86.4	<b>86.54</b>	86.0	86.4
	G	88.8	95.33	95.28	94.88	95.06	95.19	<b>95.45</b>	95.05
	B	90.85	97.53	<b>97.96</b>	97.93	97.85	97.89	97.95	97.88
	Moy	89.06	95.66	95.63	95.28	95.41	95.49	95.57	<b>95.88</b>
GM	R	82.76	83.81	84.06	84.99	87.67	84.49	84.7	<b>91.99</b>
	G	89.27	92.76	94.85	95.05	95.79	94.7	94.33	<b>95.94</b>
	B	92.47	95.69	97.53	<b>97.73</b>	97.63	97.39	97.42	97.28
	Moy	89.03	92.92	94.67	95.3	95.81	95.05	95.03	<b>96.07</b>
GZ	R	83.03	83.86	84.90	85.58	86.34	84.74	84.33	<b>87.79</b>
	G	92.47	96.30	97.8	97.86	<b>97.85</b>	94.26	94.75	96.01
	B	87.97	95.3	96.75	96.79	97.85	97.84	97.84	<b>97.89</b>
	Moy	89.47	93.29	95.31	95.06	95.41	94.81	95.1	<b>96.2</b>
H1M	R	82.74	83.08	88.08	<b>88.15</b>	87.67	87.56	87.66	86.24
	G	88.9	88.95	93.07	94.36	<b>95.79</b>	92.53	93.16	90.53
	B	93.11	92.11	96.61	<b>97.67</b>	97.63	96.29	96.42	93.86
	Moy	89.19	88.87	93.69	94.84	<b>95.81</b>	92.48	93.57	90.76
H1Z	R	82.74	84.42	88.89	88.26	86.09	88.75	88.88	<b>89.88</b>
	G	88.5	89.49	94.35	94.16	95.06	93.85	93.93	<b>95.07</b>
	B	93.05	91.84	97.69	<b>97.96</b>	97.84	97.54	97.41	97.88
	Moy	89.19	90.44	94.94	94.75	95.41	94.21	94.54	<b>95.28</b>
H2M	R	84.93	<b>92.7</b>	92.66	89.48	87.15	87.32	87.57	91.28
	G	86.56	93.52	95.89	95.59	95.79	95.54	95.81	<b>96.94</b>
	B	87.51	95.53	97.79	97.78	97.63	97.70	97.66	<b>97.85</b>
	Moy	86.48	93.48	95.89	96.12	95.81	95.65	95.82	<b>97.02</b>
H2Z	R	80.2	83.01	82.73	81.55	80.82	86.42	86.2	<b>86.53</b>
	G	81.06	91.79	91.74	91.82	91.87	95.08	95.00	<b>95.39</b>
	B	82.24	96.65	97.8	97.91	97.85	97.89	97.86	<b>97.92</b>
	Moy	81.06	92.01	91.79	91.97	92.21	95.44	95.23	<b>95.72</b>
H5M	R	<b>88.92</b>	88.01	86.72	86.83	87.67	86.77	87.04	88.34
	G	95.05	95.41	95.48	95.67	95.79	95.39	95.03	<b>96.44</b>
	B	96.95	97.16	97.6	97.72	97.63	97.53	97.53	<b>97.76</b>
	Moy	95.15	95.38	95.57	95.75	95.81	95.48	95.25	<b>96.57</b>
H5Z	R	<b>87.94</b>	86.12	85.72	84.62	86.4	85.76	85.76	86.69
	G	95.05	94.2	94.42	94.57	95.06	94.8	94.96	<b>95.48</b>
	B	96.95	97.46	97.77	97.8	97.85	97.81	<b>97.87</b>	97.85
	Moy	95.15	94.6	95.1	95.02	95.42	95.15	95.18	<b>95.85</b>

TABLE 4.2 – Résultats de classification binaire (" ciel "/" non ciel ") en fonction de la méthode de combinaison couleur/texture utilisée (fixe, hybride et proposée) et des couples gradients couleur/texture, sur la base d'images du projet CAPLOC.

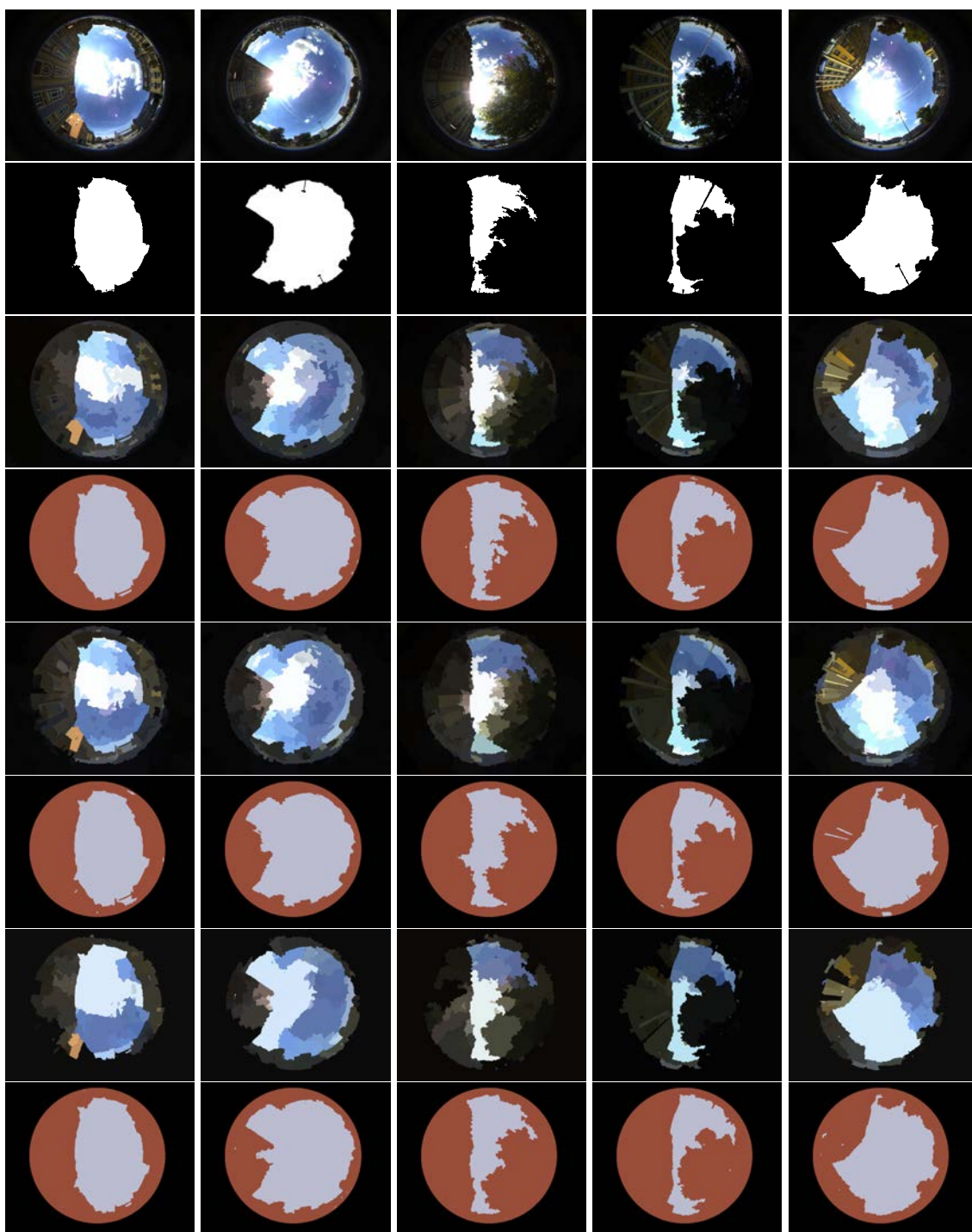


FIGURE 4.7 – Résultats de segmentations et de classifications obtenues avec l'algorithme Fisher pour chaque combinaison couleur/texture sur une sélection d'images de la base du projet CAPLOC (de haut en bas : images originales, vérités-terrain, images segmentées puis classées par la combinaison fixe, par la combinaison hybride et par la combinaison proposée).

D'après le tableau 4.2, nous constatons que les meilleurs résultats sont, dans la majorité des cas, obtenus par la combinaison couleur/texture basée sur l'ACP. En effet, celle-ci se positionne 23 fois sur 40 en première position (contre 12 fois sur 40 pour la combinaison fixe et 5 fois sur 40 pour la combinaison hybride). Nous constatons également que le meilleur taux de bonne classification, qui approche les 98%, est obtenu à chaque fois sur la composante bleue de l'espace couleur ( $R, G, B$ ).

La figure 4.7 illustre les résultats de segmentation (par combinaison couleur/texture fixe, hybride et proposée) et de classification (par l'algorithme Fisher) sur cinq images de la base CAPLOC. Pour une meilleure visualisation, seuls les résultats avec les paramètres optimaux (le paramètre de pondération et le gradient couleur/texture) sont présentés. Nous pouvons constater que les résultats obtenus par la segmentation basée sur la combinaison couleur/texture proposée sont légèrement meilleurs que les résultats obtenus avec les combinaisons fixe et hybride, même s'ils restent très proches.

Bien que la première approche de détection de " ciel "/" non ciel " proposée donne d'excellents résultats (TBC moyen proche de 98%), elle ne respecte pas la contrainte temps-réel imposée par l'application. À titre indicatif, le temps de calcul d'un gradient hybride avec la combinaison couleur/texture basée sur l'ACP est de 3 minutes et 40 secondes sur une image de la base LOCOPROL (images de taille  $288 \times 384$ ) avec un PC Core (TM2) cadencé à  $2.80GHz$ . En effet, les outils utilisés pour le calcul d'attributs de texture sont souvent coûteux en temps de calcul. Nous proposons donc, dans la suite de ce chapitre, une deuxième approche, moins coûteuse en temps de traitement, basée sur une simplification de l'image suivie d'une classification pixellaire.

#### 4.5.2/ DEUXIÈME APPROCHE : SIMPLIFICATION DE L'IMAGE ET CLASSIFICATION PIXELLAIRE

L'acquisition des bases d'images a été effectuée avec une plateforme mobile (véhicule instrumenté), ce qui a entraîné de brusques changements de luminosité. Les images présentent donc des pics brillants dus à cette variation de luminosité. En effet, la caméra fixée sur le toit du véhicule en mouvement acquiert le ciel en présence de nuages et de rayons solaires avec une forte luminosité. Par conséquent, les images peuvent présenter des niveaux de gris saturés et des régions locales homogènes et brillantes (cf. figure 4.8).



FIGURE 4.8 – Illustration de la présence de nuages et de forte illumination du soleil sur une sélection d'images issues de la base CAPLOC.

Ces phénomènes (nuages et variation de luminosité) rendent la classification d'une image en deux classes (" ciel "/" non ciel ") plus complexe. Il paraît donc opportun



d'homogénéiser l'image à l'aide d'une étape de lissage/simplification avant de classer les pixels de l'image afin de supprimer les pics locaux brillants.

Dans ce qui suit, nous rappelons la définition de quelques filtres les plus communément utilisés dans la littérature. Ces filtres seront testés sur les deux bases d'images LOCOPROL et CAPLOC. Nous reprenons ci-après la classification des filtres faite par MANZANERA [Manzanera] à savoir une répartition en deux catégories : les filtres de lissage et les filtres dérivateurs (cf. figure 4.9).

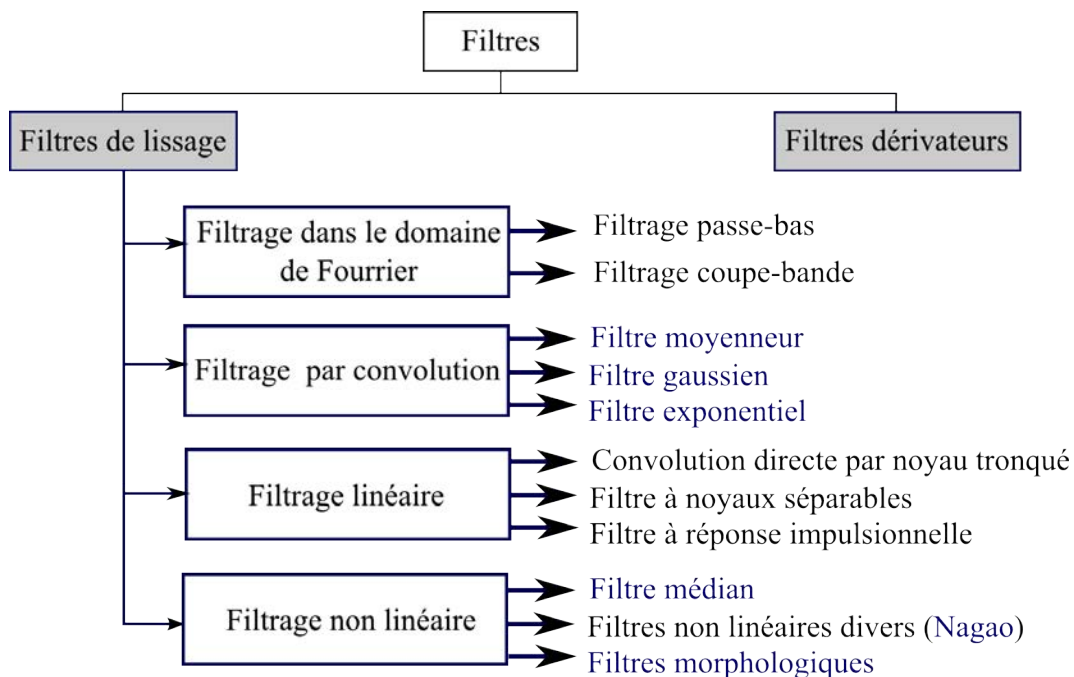


FIGURE 4.9 – Les différents types de filtre dans la littérature.

Dans notre cas, nous visons le lissage de l'image en éliminant les pics brillants présents, tout en gardant les détails et les contours des objets. Nous avons alors fait recours à la première catégorie de filtres (filtres de lissage), qui est à son tour divisée en quatre sous-catégories :

- **filtres définis dans le domaine de Fourier** tels que les filtres passe-bas et les filtres coupe-bande ;
- **filtres par convolution** tels que le filtre moyennneur, le filtre gaussien et le filtre exponentiel ;
- **filtres linéaires** tels que les filtres issus de la convolution directe par noyau (tronqué), les noyaux séparables et les filtres à réponse impulsionnelle ;
- **filtres non linéaires** qui se caractérisent par l'élimination du bruit impulsionnel dans l'image et par la conservation des contours des objets dans l'image. Nous pouvons citer à titre d'exemple le filtre médian, les filtres morphologiques (filtres alternés séquentiel, filtre de nivellement, les reconstructions géodésiques, etc.)

Compte tenu de l'application visée, nous testerons seulement, dans ce qui suit, les filtres les plus rapides pour tenir compte de la contrainte temps-réel.



#### 4.5.2.1/ FILTRE MÉDIAN

Le filtre médian consiste en une opération de lissage non linéaire en parcourant l'image et en remplaçant la valeur du pixel courant par la valeur médiane de ses voisins (définis dans un voisinage de taille  $W$ ), comme illustre l'exemple dans la figure 4.10. Le filtre médian est caractérisé par sa capacité à lisser tout bruit impulsionnel ou exponentiel, ce qui paraît adapté à notre objectif. Notons que plus la taille du voisinage est grande, plus le lissage est fort et donc plus les détails de l'image sont atténués.

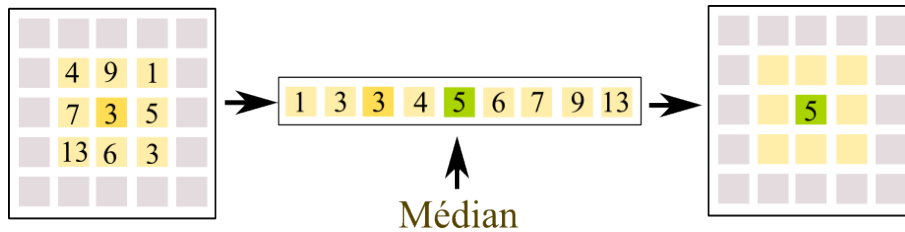


FIGURE 4.10 – Illustration du principe du filtre médian avec  $W = 3$ .

La figure 4.11 illustre sur une image de la base LOCOPROL l'effet du filtre médian en fonction de la taille du voisinage  $W$ .



FIGURE 4.11 – Illustration du lissage par filtre médian avec différentes valeurs de voisinage  $W$ .

#### 4.5.2.2/ FILTRE MOYENNEUR

Le filtre moyennneur est un filtre linéaire spatial isotrope basé sur le principe que la valeur d'un pixel est relativement similaire à son voisinage. Ce filtre consiste à appliquer un masque de convolution de taille  $W$ , puis remplacer la valeur du pixel courant par la moyenne pondérée de ses voisins. Les coefficients du masque représentent les coefficients de pondération. L'équation ci-dessous présente un exemple d'un masque de convolution  $h$  de taille  $W = 3$ .

$$h = \frac{1}{9} \times \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (4.6)$$

La figure 4.12 illustre, sur une image issue de la base LOCOPROL, l'effet du filtre moyennneur en fonction du paramètre  $W$ . Nous pouvons noter que le lissage préserve toutes les zones de l'image ayant un niveau de gris constant (zones homogènes), mais que l'image devient de plus en plus floue avec l'augmentation de la taille du masque de convolution. En effet, sur la cinquième image de la figure 4.12 (ie.  $W = 8$ ), les contours deviennent très flous, contrairement à ceux issus du filtre médian pour une même taille

du masque de convolution. De plus, le filtre moyennneur a l'inconvénient d'engendrer des phénomènes de fausses couleurs, contrairement au filtre médian.

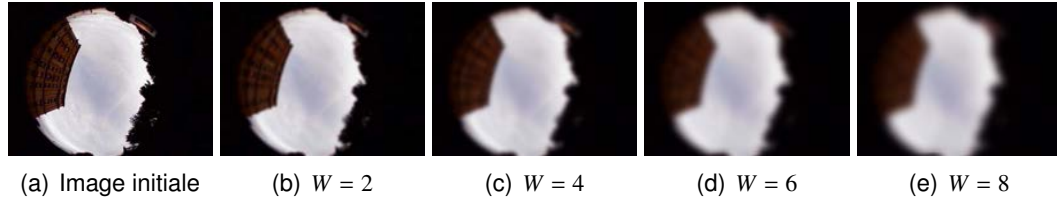


FIGURE 4.12 – Illustration du lissage par filtre moyennneur avec différentes valeurs de voisinage  $W$ .

#### 4.5.2.3/ FILTRE GAUSSIEN

Un lissage par filtre gaussien est un lissage par convolution caractérisé par le paramètre  $\sigma$ , correspondant à la largeur du filtre et l'écart type de la fonction gaussienne utilisée et donnée par l'équation 4.7. Ce paramètre prend ses valeurs dans l'intervalle  $[0, TailleImage/6]$ . Cette opération de lissage est similaire dans son principe au filtre moyennneur, avec un masque de convolution ayant des poids plus élevés au centre.

$$g(i, j) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right) \quad (4.7)$$

La figure 4.13 illustre, sur une image issue de la base LOCOPROL, l'effet du filtre gaussien en fonction du paramètre  $\sigma$ . Notons que plus la valeur de  $\sigma$  est élevée (ie. plus la "cloche" de la gaussienne est large), plus l'image est lissée (ie. floue).

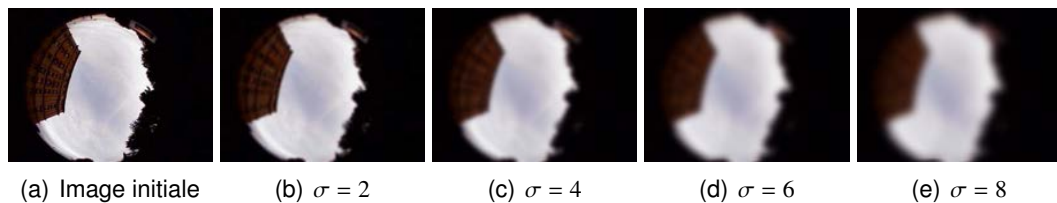


FIGURE 4.13 – Illustration du lissage par filtre gaussien avec différentes valeurs de  $\sigma$ .

#### 4.5.2.4/ FILTRE EXPONENTIEL SYMÉTRIQUE

Le filtre exponentiel synthétique est réalisé à l'aide de deux fonctions, données par les équations 4.8 et 4.9. La première fonction est appliquée, dans un premier temps, et le résultat est fourni en entrée de la deuxième fonction. Le filtre exponentiel utilise un paramètre, noté  $\alpha$ , permettant de faire varier le degré de lissage d'une image  $I$ . Plus ce paramètre est faible, plus le lissage est fort (cf.figure 4.14).

$$e_1(i) = \alpha \times (I(j, i) - e_1(i - 1)) + e_1(i - 1) \quad (4.8)$$

$$e_2(i) = \alpha \times (e_1(i) - e_2(i - 1)) + e_2(i - 1) \quad (4.9)$$

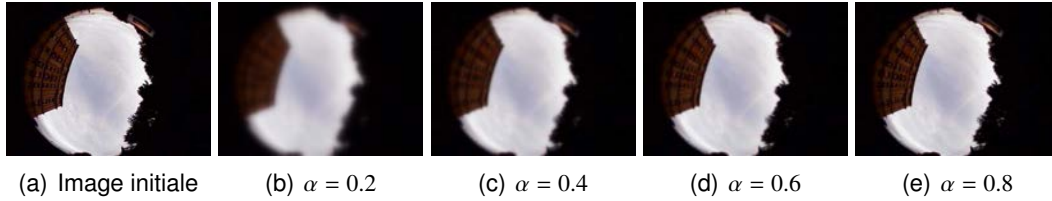


FIGURE 4.14 – Illustration du lissage par filtre exponentiel avec différentes valeurs de  $\alpha$ .

#### 4.5.2.5/ FILTRE DE MALIK-PERONNA

L'algorithme de MALICK-PERONA [Perona 1990] consiste à lisser l'image par une diffusion non linéaire, en utilisant des dérivées finies de premier ordre. Ceci permet de lisser les zones homogènes de l'image en préservant les contours des objets. Le filtre correspondant est donné par l'équation 4.10.

$$\begin{cases} \frac{\partial I}{\partial itr} = \text{div}(c(|\nabla I|)\nabla I) \\ I(i, j, itr = 0) = I_0(i, j) \end{cases} \quad (4.10)$$

où  $c$  est la fonction qui conditionne la diffusion de sorte que cette dernière soit forte dans les zones à faibles gradients et faible dans les zones à forts gradients. Cette fonction d'arrêt peut être formulée par les équations 4.11 ou 4.12.

$$c(I) = \exp\left(-\left(\frac{|\nabla I|^2}{h_c}\right)\right) \quad (4.11)$$

ou

$$c(I) = \frac{1}{1 + \left(\frac{|\nabla I|^2}{h_c}\right)} \quad (4.12)$$

Le paramètre  $h_c$  correspond à la hauteur des contours à préserver. Il influence, avec le paramètre du nombre d'itérations  $itr$ , le résultat du lissage. La figure 4.15 illustre, sur une image issue de la base du projet LOCOPROL, l'évolution du résultat de lissage par le filtre de MALIK-PERONA en fonction du nombre d'itérations  $itr$  avec  $h_c = 10$ .

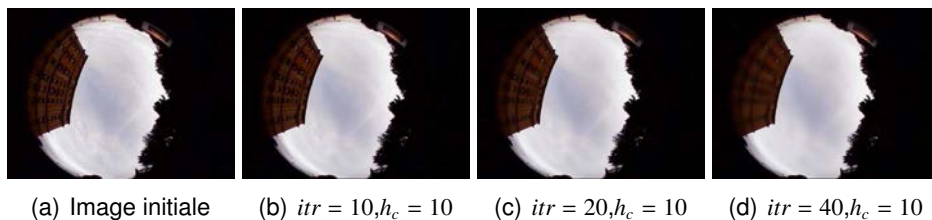


FIGURE 4.15 – Illustration de l'évolution du lissage par filtre de MALIK-PERONA en fonction du nombre d'itérations  $itr$ .

#### 4.5.2.6/ FILTRE DE NAGAO

Le lissage de NAGAO est basé sur le fractionnement du voisinage du pixel courant en différents domaines. En effet, pour chaque pixel  $p$ , un voisinage de taille  $5 \times 5$  est considéré et 9 domaines distincts sont par la suite définis (cf. figure 4.16). Pour chaque domaine  $D_p$ , la moyenne  $\mu_p$  et la variance  $v_p$  des niveaux de gris sont calculées. Le domaine le plus homogène et ayant la variance la plus faible est finalement retenu. La valeur du pixel d'intérêt (le point central de la fenêtre  $5 \times 5$ ) est remplacée par la valeur moyenne des niveaux de gris du domaine sélectionné.

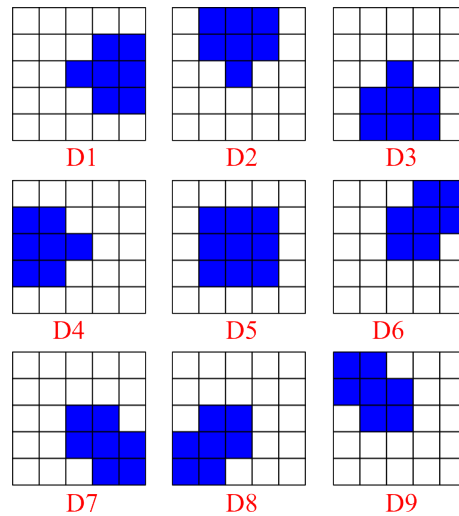


FIGURE 4.16 – Les 9 domaines de NAGAO.

L'avantage du filtre de NAGAO est qu'il permet de réduire le bruit tout en renforçant les contrastes par diminution de la largeur de bande de la zone de transition. La figure 4.17 illustre, sur une image issue de la base LOCOPROL, l'effet du filtre de NAGAO.

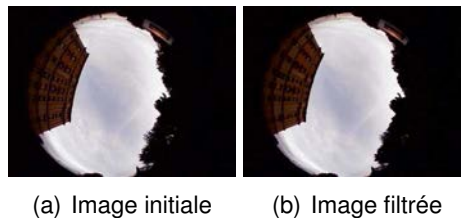


FIGURE 4.17 – Illustration du lissage par filtre de NAGAO.

#### 4.5.2.7/ FILTRE ADAPTATIF BASÉ SUR LE VOISINAGE

Un lissage par filtre adaptatif basé sur le voisinage (OutRange) revient à appliquer un lissage linéaire en conservant les points de fort contraste. Cette opération consiste à remplacer la valeur du pixel courant, par la valeur moyenne de ses voisins, à l'issue d'une comparaison par rapport à un seuil fixé a priori, de l'écart type entre cette moyenne et la valeur du pixel central (cf. figure 4.18). Notons qu'une valeur nulle de ce seuil correspond à un lissage par filtre moyennneur. La figure 4.19 illustre le fait que plus l'écart entre la

valeur du pixel central et la moyenne de ses voisins est élevée, plus l'effet de flou est important.

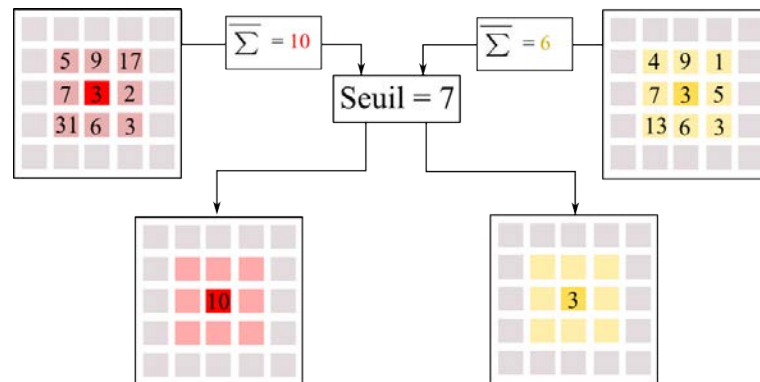


FIGURE 4.18 – Principe du filtre adaptatif basé sur le voisinage.



FIGURE 4.19 – Illustration du lissage par filtre adaptatif basé sur le voisinage avec différentes valeur du seuil.

#### 4.5.2.8/ FILTRE ADAPTATIF SYMETRIC NEAREST NEIGHBOURHOOD (SNN)

Le filtre adaptatif Symetric Nearest Neighbourhood (SNN) permet de réduire le bruit dans l'image tout en préservant les contours.

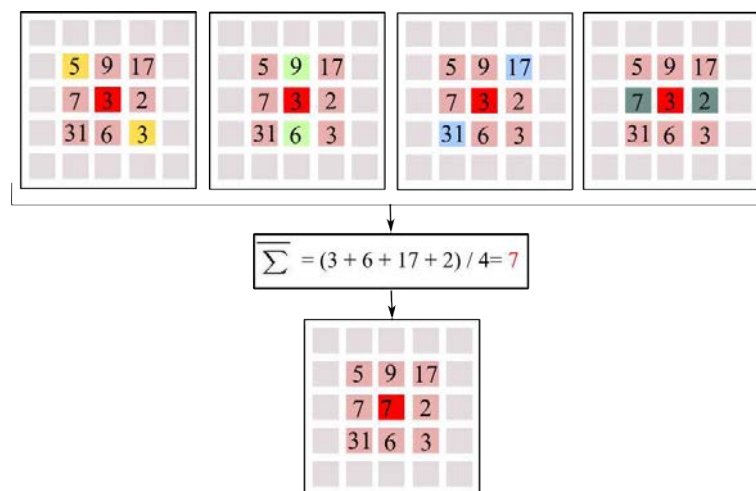


FIGURE 4.20 – Principe de calcul du filtre SNN.

Le principe de ce filtre consiste à comparer les paires de pixels voisins opposés dans un voisinage déterminé. Ensuite, les pixels dont la valeur est plus proche de celle du pixel central sont retenus. Finalement, la valeur du pixel d'intérêt (pixel central) est remplacée par la moyenne des valeurs des pixels retenus. La figure 4.20 illustre le principe du filtre SNN, avec un voisinage  $3 \times 3$ . Ainsi, quatre paires de pixels sont à comparer.

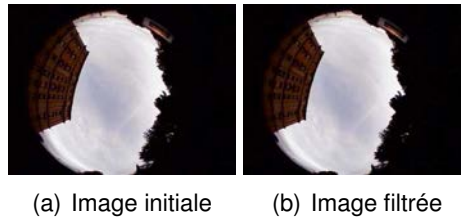


FIGURE 4.21 – Illustration du lissage par filtre adaptatif Symetric Nearest Neighbourhood.

#### 4.5.2.9/ RECONSTRUCTION GÉODÉSIQUE PAR DILATATION (RGD)

Pour décrire précisément l'opérateur de reconstruction géodésique par dilatation (RGD), il est nécessaire de rappeler quelques définitions basiques de morphologie mathématique. Ces définitions sont présentées avec beaucoup plus de détails dans les travaux [Serra 1988, Shmitt 1994].

**Définition 1** : Un maximum régional, noté  $M$ , d'une image  $I$  est un plateau tel que les niveaux de gris des pixels voisins sont inférieurs à la valeur  $I(M)$ .

**Définition 2** : Un plateau en un pixel  $p$  d'une image  $I$ , noté  $Plt_p(I)$ , est une composante connectée contenant  $p$  et ayant un niveau de gris constant égal à  $I(p)$ .

Notons par  $Max_{reg}(I)$  l'ensemble des maximums régionaux de l'image  $I$ . Un maximum régional,  $M$  peut être un seul pixel, mais il peut être une région de l'image. Par conséquent, un maximum régional  $M \subseteq IR^2$ . Formellement,  $M$  est un maximum régional si et seulement si, il existe un voisinage  $\mathcal{V}_M$  qui satisfait l'équation 4.13.

$$\forall p \in \mathcal{V}_M \setminus \{M\} \quad I(p) < I(M) \quad (4.13)$$

La reconstruction géodésique par dilatation est une technique des plus utilisées pour l'extraction des maximums régionaux [Sbeh 2001]. En effet, cette technique fait partie des méthodes de filtrage utilisant des opérateurs de morphologie mathématique (dilatation, érosion, ouverture, fermeture, ASF, etc). La RGD utilise deux images en entrée : une image *masque*  $I$  et une image *marqueur*  $I'$ , vérifiant la condition  $\forall p, I'(p) < I(p)$ . Le principe de la RGD visant à extraire des maximums régionaux de l'image consiste à appliquer une suite finie de dilatation géodésique jusqu'à idempotence.

**Définition 3** : L'idempotence indique la stabilité du résultat obtenu. Elle est atteinte



lorsque le résultat obtenu à l'itération  $(n + 1)$  est identique celui obtenu à l'itération  $n$ .

Dans le cas usuel, la formule de la RGD est donnée par l'équation 4.16 [Vincent 1993] :

$$I_{RGD} = \bigvee_{(n \geq 1)} \delta_I^{(n)}(I') \quad (4.14)$$

$$\delta_I^{(n)}(I') = \delta_I^{(1)} \circ \delta_I^{(1)} \circ \dots \circ \delta_I^{(1)}(I') \quad (4.15)$$

$$\delta_I^{(1)}(I') = (I' \oplus B) \wedge I \quad (4.16)$$

avec

$$\forall p \in I, \quad I'(p) < I(p) \quad (4.17)$$

où  $\oplus$  correspond à une dilatation morphologique avec l'élément structurant  $B$  et  $\delta_I^{(1)}(I')$  est le premier rang de dilatation géodésique.  $\delta_I^{(n)}(I')$  correspond au résultat de composition (notée par  $\circ$ ) des dilatations géodésiques jusqu'à stabilité (idempotence donnée par le  $n^{eme}$  rang).

À partir de l'équation 4.16 et en considérant  $\forall p \in I, I'(p) = I(p) - 1$ , nous pouvons obtenir l'ensemble des maximums régionaux  $Max_{reg}(I)$  donné par l'équation 4.18 ci-après :

$$Max_{reg}(I) = \{p \in I / (I - I_{RGD})(p) > 0\} \quad (4.18)$$

Dans le cas d'images naturelles, les maximums régionaux ont généralement une différence de niveaux de gris supérieure à 1. Soit  $H$  la hauteur d'un maximum régional (appelé aussi un pic), l'image *marqueur*  $I'$  est obtenue en soustrayant à  $I$  la valeur de  $H : \forall p \in I, I'(p) = I(p) - H$ .

La figure 4.22 illustre les différentes étapes de la reconstruction géodésique par dilatation sur une ligne (fonction 1D) d'une image (fonction 2D). Nous pouvons noter aisément que la RGD a permis de supprimer les pics locaux de la ligne et d'obtenir alors des zones plus homogènes (en niveaux de gris).

Pour une meilleure visualisation, la figure 4.23 illustre en 2D puis en 3D, l'effet du lissage de la zone ciel sur une image issue de la base CAPLOC. Nous pouvons constater à partir des deux représentations 3D de l'intensité de l'image initiale (à gauche) et de l'image simplifiée (à droite), que les variations entre les maximums locaux de l'image s'atténuent et que les régions de l'image deviennent plus homogènes (suppression des pics de l'image).

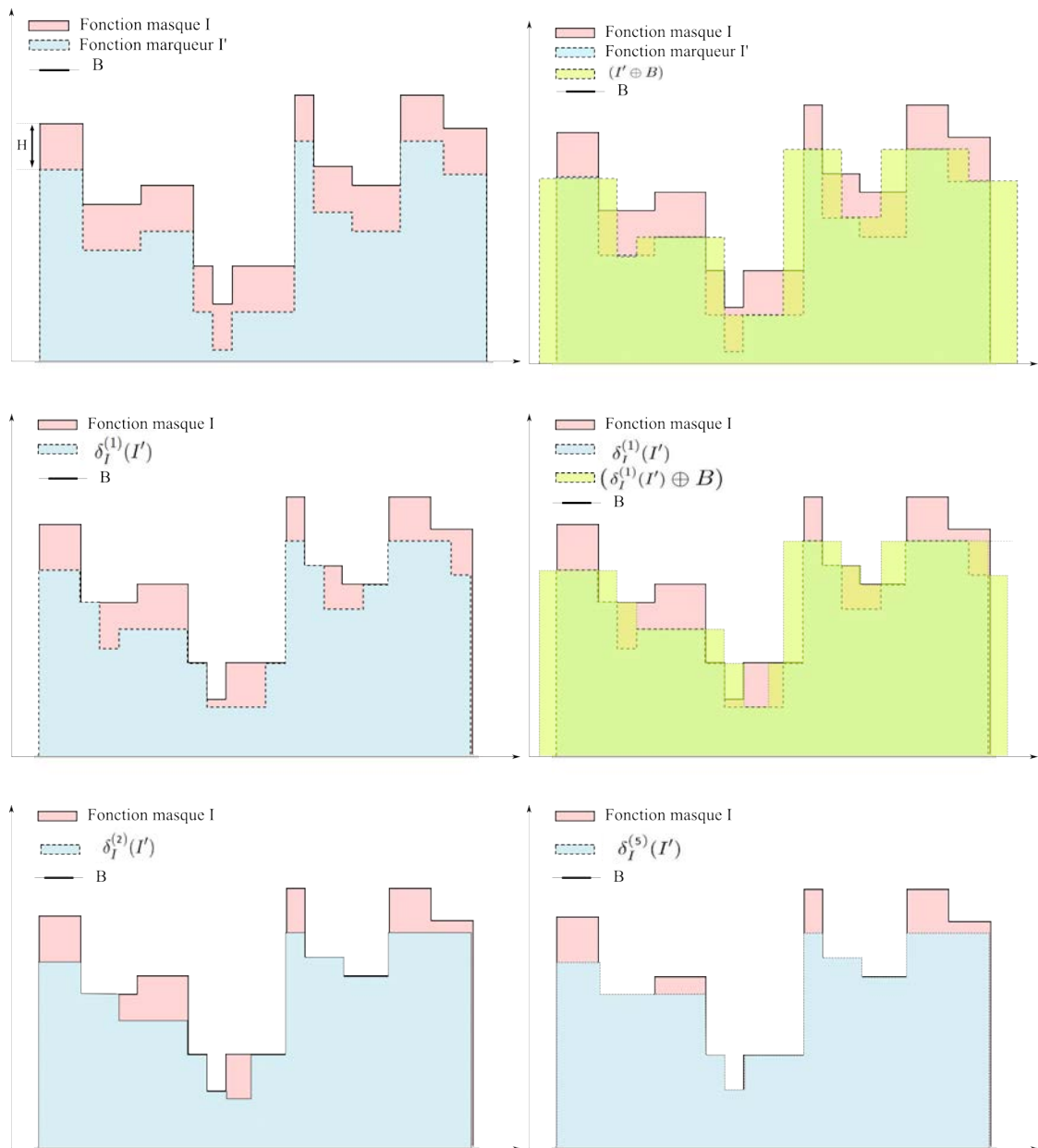


FIGURE 4.22 – Effet de la RGD sur une fonction 1D.

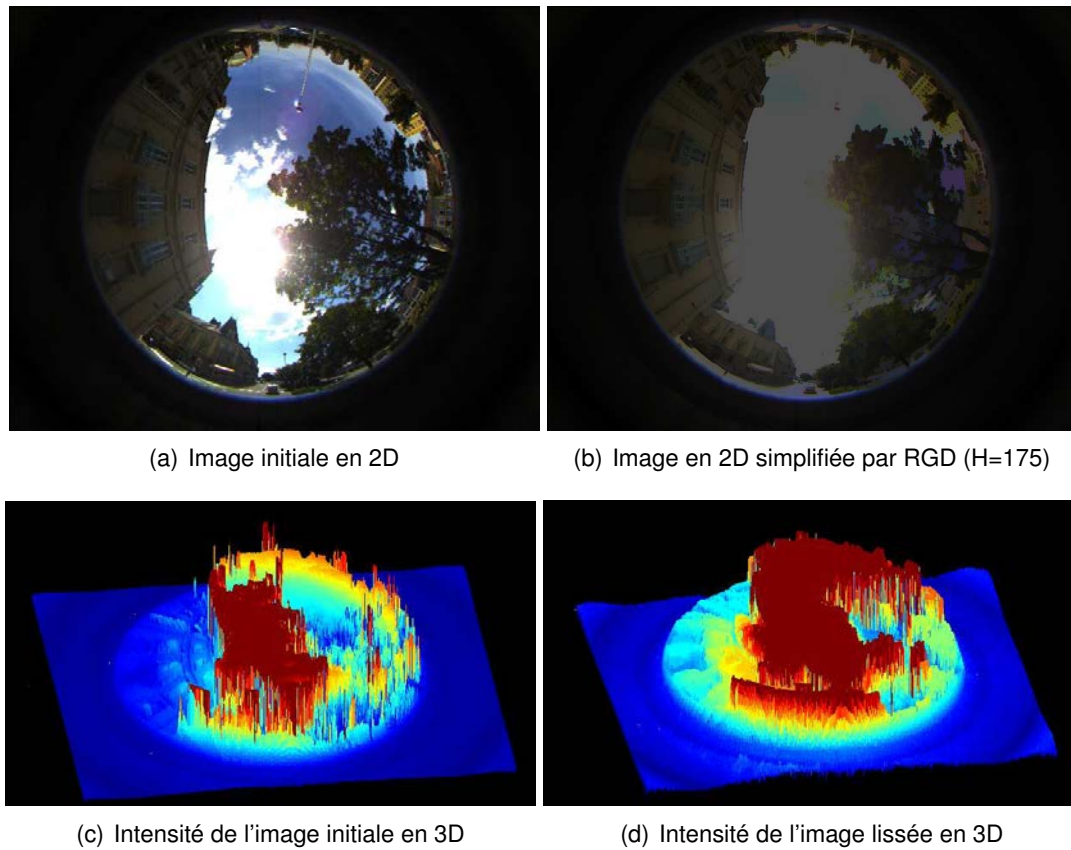


FIGURE 4.23 – Illustration de l'effet de la simplification par RGD sur une image issue de la base CAPLOC, avec la représentation 3D de l'intensité.

Il convient également de noter que la simplification de l'image par RGD varie en fonction du paramètre de contraste  $H$ , comme l'illustre la figure 4.24. Ce paramètre doit ainsi influencer le taux de bonne classification final. Comme le montre la figure (4.24), en faisant varier la valeur du paramètre de contraste  $H$ , le résultat de la RGD change, et par conséquent, le taux de bonne classification change.

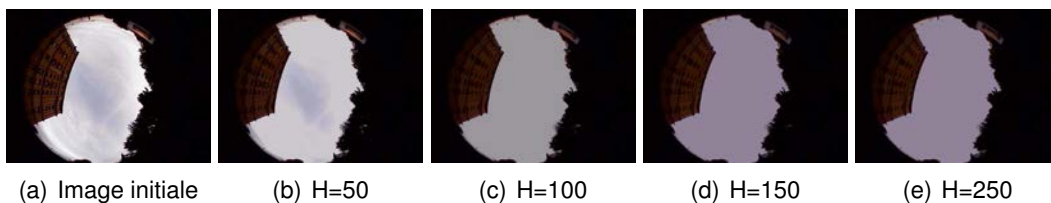


FIGURE 4.24 – Influence de la valeur du paramètre de contraste  $H$  sur la simplification d'image par RGD.

Disposant maintenant d'une image simplifiée, nous donnons, dans ce qui suit, la définition de quelques algorithmes de classification (non supervisés et supervisés) qui seront utilisés et comparés, pour classer les pixels de l'image fisheye en deux classes : " ciel " et " non ciel ".

### 4.5.3/ CLASSIFICATION PIXELLAIRE

La problématique de la classification se pose dans différents domaines comme l'extraction de données [Bolshakova 2003, Osei-Bryson 2010], la compression de données et la quantification vectorielle [Yanga 2008, Lughofer 2008], la reconnaissance et la classification de formes [Kyperountasa 2008, Lo 2012], etc. Les méthodes de classification sont généralement classées en différentes familles comme les méthodes basées sur la décomposition/fusion (ou split/merge en anglais) [Nock 2004b], les méthodes aléatoires [Moosmann 2008], les méthodes basées sur les réseaux de neurones [Gupta 2008], les méthodes dédiées aux bases de données de grandes dimensions [Lafarge 2004], etc. Dans ce chapitre, nous préférons adopter une catégorisation plus globale des algorithmes de classification à savoir les algorithmes de classification non-supervisée et les algorithmes de classification supervisée. Nous présentons ci-après, les algorithmes les plus communément utilisés dans la littérature, figurant parmi ces deux catégories. Le lecteur intéressé par un nombre plus élevé d'algorithmes ou de précisions sur ces derniers pourra se reporter aux travaux référencés ci-après : [Kanungo 2003, Mount 2005, Nielsen 2003, Nock 2004a, Lefèvre 2009, Mathankera 2011].

#### 4.5.3.1/ MÉTHODES DE CLASSIFICATION NON-SUPERVISÉE

Dans cette section, nous rappelons brièvement le principe de quelques algorithmes de classification non supervisés, à savoir l'algorithme KMLocal (présentés avec ses différentes versions), l'algorithme des C-moyennes floues (ou en anglais le Fuzzy Cmeans), l'algorithme de Fisher et enfin l'algorithme Statistical Region Merging (SRM).

##### a) L'algorithme KMLocal

L'algorithme de classification KMLocal reprend les différentes versions d'implémentations possibles de l'algorithme de classification des k-moyennes. En effet, ces implémentations ont été développées afin de réduire le temps attribué à la recherche du plus proche centre de gravité (dans l'algorithme classique k-moyennes). Cet algorithme regroupe quatre versions améliorées : Lloyd's, Swap, Hybrid et EZ-Hybrid. Celles-ci sont présentées ci-après.

Nous commençons tout d'abord par rappeler la définition de base de l'algorithme des k-moyennes. La classification par l'algorithme des k-moyennes a été proposée par MACQUEEN [MacQueen 1967] et consiste à regrouper un ensemble de points initial  $P$  formé par  $N$  points, de dimension  $d$  chacun, en  $C$  sous ensembles (que nous appellerons par la suite "classes") :

$$P = \{p_n\}_{n=1\dots N} \quad (4.19)$$

où

$$p_n \in \mathbb{R}^d, \forall n = 1 \dots N$$

Chaque classe  $C_i$  possède un centre  $w_i$  qui correspond concrètement à son centre de gravité. L'algorithme des k-moyennes consiste à minimiser la distance quadratique moyenne entre chaque élément  $p_n$  appartenant à un sous ensemble  $\{C_i\}_{1 \leq i \leq C}$  et son

centre  $w_i$ , afin d'associer cet élément à la classe dont le centre de gravité est le plus proche. La recherche du centre le plus proche est l'étape la plus complexe face à la difficulté à trouver un minimum global. Cette complexité est résolue par une autre version de l'algorithme des k-moyennes, connue sous le nom de Lloyd's, qui vise à chercher les minimums locaux de l'équation 4.20. Cette version a été introduite par [Lloyd 1982] et est utilisée par plusieurs auteurs [Selim 1984, Kanungo 2003].

$$\operatorname{argmin}_{1 \leq i \leq C} \|p_n - w_i\|^2 \quad (4.20)$$

Dans ce qui suit, nous détaillons les différentes versions de l'algorithme KMlocal :

- **Lloyd's** est souvent confondu avec l'algorithme des k-moyennes. Il se base sur un processus itératif dont le but est de trouver les  $C$  centres donnant la meilleure classification possible. Lloyd's commence par sélectionner  $C$  centres initiaux choisis aléatoirement parmi les  $N$  points. Il répète ensuite la recherche des voisins les plus proches à chaque centre  $w_i$ , ce qui permet de former les sous-ensembles initiaux. L'étape suivante consiste à chercher les nouveaux centres de gravité. Ces deux étapes sont répétées jusqu'à atteindre un nombre maximal d'itérations fixé a priori ou valider par une condition de convergence, comme par exemple la stabilité des centres de gravité. La qualité de la classification dépend du premier choix intuitif des centres de gravité et du nombre maximal d'itérations.
- **Swap** est un algorithme qui intervient au niveau de la phase de changement des centres. En effet, cet algorithme sélectionne initialement un sous ensemble de  $C$  centres initiaux parmi les centres candidats. Il vise ensuite, l'échange des centres, un par un, en comparant le résultat de la classification selon un critère mesurant la distorsion. Soient  $\Omega$  l'ensemble des centres candidats et  $\Psi$  un sous-ensemble de  $\Omega$ . Swap remplace un centre  $\psi \in \Psi$  par un nouveau candidat  $\psi' \in \Omega - \Psi$  afin d'obtenir un nouveau sous-ensemble de centres  $\Psi' = \Psi - \{\psi\} \cup \{\psi'\}$ . Le sous-ensemble  $\Psi$  est remplacé par  $\Psi'$  si la nouvelle solution réduit le critère de distorsion. L'algorithme converge quand le critère de distorsion est inchangé.
- **Hybrid** est un algorithme qui combine les deux algorithmes Lloyd's et Swap avec des appariements variés.
- **EZ-Hybrid** est une version hybride moins complexe que la version hybride seule. EZ-Hybrid applique, une première fois, la version Swap suivie d'un ensemble d'itérations de l'algorithme Lloyd's. Ce dernier est appliqué jusqu'à obtenir la RDL (Relative Distorsion Loss), donnée par l'équation 4.21, inférieure à un seuil donné.

$$RDL = \frac{\text{DistorsionInitiale} - \text{DistorsionCourante}}{\text{DistorsionInitiale}} \quad (4.21)$$

#### b) L'algorithme Fuzzy Cmeans

L'algorithme Fuzzy Cmeans (ou C-moyennes floues en français) a été développé par [Dunn 1973], puis amélioré par [Bezdek 1981]. Il est proche dans son principe de l'algorithme des k-moyennes. Il vise à classer un ensemble de  $N$  points en  $C$  classes. Cet algorithme est basé sur la minimisation d'un critère  $J_m$  qui est la somme de la distance quadratique pondérée (par des coefficients d'appartenance  $\mu_{ni}$ ) d'un point  $p_n$  ( $1 \leq n \leq N$ ) à une classe  $C_i$ . Le critère  $J_m$  est donné par l'équation (4.22) où le

paramètre  $m \geq 1$  désigne le degré de flou.

$$J_m = \sum_{n=1}^N \sum_{i=1}^C \mu_{ni}^m \|p_n - w_i\|^2 \quad (4.22)$$

Les coefficients d'appartenance  $\mu_{ni}$  forment la matrice d'appartenance notée  $U$  et donnée par l'équation (4.23). Chaque terme de cette matrice d'appartenance est défini par l'équation (4.24) et possède une valeur comprise dans l'intervalle  $[0, 1]$ .

$$U_{N \times C} = \{\mu_{ni}^m\}_{\substack{n=1 \dots N \\ i=1 \dots C}} \quad (4.23)$$

$$\mu_{ni}^m = \frac{1}{\sum_{j=1}^C \left( \frac{\|p_n - w_i\|}{\|p_n - w_j\|} \right)^{\frac{2}{m-1}}} \quad (4.24)$$

avec  $w_i$  donné par l'équation (4.25)

$$w_i = \frac{\sum_{n=1}^N \mu_{ni}^m \cdot p_n}{\sum_{n=1}^N \mu_{ni}^m} \quad (4.25)$$

L'algorithme des C-moyennes floues débute par une initialisation de la matrice d'appartenance  $U^{(0)}$ . La deuxième étape consiste à calculer, à l'aide de l'équation 4.25, les centres  $w_i$ . Dans la troisième étape, la nouvelle matrice d'appartenance  $U^{(j+1)}$  est calculée à partir de  $U^{(j)}$  en utilisant l'équation (4.24). La condition d'arrêt est définie par l'inégalité donnée par l'équation 4.26.

$$\|U^{(j+1)} - U^{(j)}\| < \varepsilon \quad (4.26)$$

### c) L'algorithme de Fisher

L'algorithme de Fisher fait partie des algorithmes de classification non supervisés et a la particularité de ne s'appliquer que sur des images en niveaux de gris [Fisher 1936, Fisher 1958]. L'idée est de diviser l'histogramme de l'image (en niveaux de gris) en  $C$  classes en minimisant la somme des variances des classes. En effet, l'algorithme de Fisher consiste à localiser des seuils représentant les séparateurs de la distribution des données. Soient  $s_i$  le seuil associé à la classe  $C_i$  et  $m_i$  la moyenne des intensités des pixels appartenant à la classe  $C_i$ . La répartition des pixels en  $C$  classes revient à localiser les seuils  $\{s_i\}_{i=1 \dots C}$  comme des séparateurs de distributions en utilisant un critère qui minimise de la somme des inerties associées aux différentes classes :

$$Histog(s_i) = \min_{p \in ]m_i, m_{i+1}[} (Histog(p)) \quad (4.27)$$

À chaque classe  $C_i$ , une taille  $T_i$  (donnée par l'équation 4.28) et une variance  $V_i$  (donnée par l'équation 4.29) sont associées.

$$T_i = \sum_{p \in D_i} \{Histog(p)\} \quad (4.28)$$

$$V_i = \frac{1}{T_i} \sum_{p \in D_i} (p - m_i)^2 \cdot Histog(p) \quad (4.29)$$



$D_i = ]s_{i-1}, s_i[$  est la gamme des niveaux de gris d'une classe  $C_i$ . Le critère de minimisation est donné par l'équation (4.30) :

$$W = \sum_{i=1}^C T_i \cdot V_i \quad (4.30)$$

où  $m_i$  représente la moyenne pondérée, donnée par l'équation (4.31)

$$m_i = \frac{1}{T_i} \cdot \sum_{p \in D_i} p \cdot \text{Histogram}(p) \quad (4.31)$$

#### d) L'algorithme Statistical Region Merging (SRM)

L'algorithme Statistical Region Merging (SRM) est une approche de segmentation/classification basée sur une nouvelle disposition des pixels dans l'image. En effet, une multitude d'approches basées sur la fusion de régions et visant la segmentation/classification de l'image existe [Zhu 1996]. L'algorithme SRM, introduit par NIELSON et NOCK [Nielsen 2003] et amélioré dans leurs travaux [Nock 2004a, Nock 2005], consiste à définir un prédicat de fusion de régions ainsi que l'ordre de fusion. L'algorithme SRM commence par définir les couples de pixels adjacents  $(p, p')$ . Soient  $S_I$  l'ensemble qui regroupe ces couples et  $I$  l'image initiale à classer. Une fonction  $f(p, p')$  est définie afin d'ordonner  $S_I$  dans l'ordre croissant. Ensuite, l'ensemble  $S_I$  est parcouru une seule fois en vérifiant un prédicat de fusion. En effet, pour chaque couple  $(p, p') \in S_I$  de labels différents  $L(p) \neq L(p')$ , le prédicat de fusion courant  $P(L(p), L(p'))$  est évalué et la fusion de régions n'est donc effectuée que, si et seulement si,  $P(L(p), L(p'))$  est vrai (cf. équation (4.32)).

$$P(L(p), L(p')) = \begin{cases} \text{vrai} & \text{si } |\bar{L}' - \bar{L}| \leq \sqrt{b^2(L) + b^2(L')} \\ \text{faux} & \text{sinon} \end{cases} \quad (4.32)$$

où  $\bar{L}$  correspond à la valeur moyenne de niveau de gris de la région  $L$  et  $b(L)$  est donné par l'équation 4.33.

$$b(L) = g \sqrt{\left(\frac{1}{(2Q|L|)}\right) \ln\left(\frac{|\mathcal{R}_L|}{\eta}\right)} \quad (4.33)$$

où  $\mathcal{R}_L$  est l'ensemble des régions ayant  $l$  pixels,  $\eta$  est une constante vérifiant  $0 < \eta \leq 1$ ,  $Q$  est le paramètre de quantification de complexité statistique de l'image classée, et  $\{1, 2, \dots, g\}$  est l'ensemble des valeurs de niveau de gris (dans la pratique  $g = 256$ ).

#### 4.5.3.2/ MÉTHODES DE CLASSIFICATION SUPERVISÉE

Dans cette section, nous présentons brièvement le principe de trois algorithmes de classification supervisés communément utilisés, à savoir l'algorithme des K-plus proches voisins (Kppv), l'algorithme de Bayes et l'algorithme des Séparateurs à Vaste Marges (SVM). Ces classifieurs dits supervisés nécessitent donc la création d'une base d'apprentissage en plus de la base de tests. En effet, une classification supervisée requiert une phase d'apprentissage (effectuée généralement hors ligne) qui consiste à calculer à partir d'un ensemble d'images dites prototypes un ensemble

d'attributs. Ces derniers servent à caractériser les classes d'intérêt. À l'aide des résultats d'apprentissage, un classifieur pixellaire détermine la classe d'appartenance de chaque pixel en fonction de ses caractéristiques. La phase d'apprentissage peut être effectuée en trois étapes, à savoir l'acquisition/sélection des images prototypes, la génération des caractéristiques de chaque classe d'intérêt et la construction des classes (cf. figure 4.25). L'apprentissage étant effectué, la classification pixellaire d'une image peut opérer en ligne. Les attributs pertinents déterminés lors de l'apprentissage sont extraits au niveau du pixel à classer et comparés ensuite aux caractéristiques discriminantes afin d'assigner ce pixel à une classe (cf. figure 4.25).

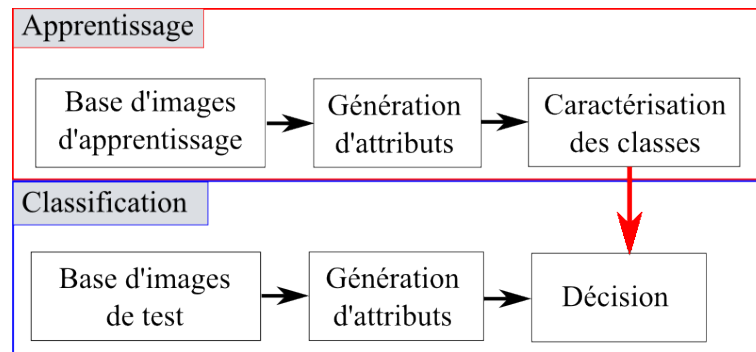


FIGURE 4.25 – La classification supervisée.

#### a) L'algorithme K-Plus proches voisins

L'algorithme de classification des K-plus proches voisins (Kppv) appartient à l'ensemble des méthodes de classification par apprentissage. Cet algorithme commence par chercher les  $K$  plus proches voisins en utilisant une mesure de distance (la distance euclidienne est souvent utilisée). Ensuite, l'occurrence de chaque classe est calculée en se basant sur les résultats d'apprentissage. La dernière étape consiste à attribuer le point courant  $p$  à la classe  $C_i$  la plus fréquente. Une première formalisation du problème a été présentée dans [Guigue 2003] sous la forme de l'équation 4.34 ci-dessous :

$$I^C(p) = C_i \text{ avec } OCC(C_i) > OCC(C_j) \quad \forall i \neq j \in \{1, 2, \dots, C\} \quad (4.34)$$

où  $OCC$  est l'opérateur définissant le nombre d'occurrences de  $C_i$  parmi les labels des  $K$  plus proches voisins de  $p$ .  $I^C$  désigne l'image classée.

Une deuxième formalisation de l'algorithme des Kppv est basé sur l'estimation d'une densité de probabilité (cf. équation 4.35) à la place d'une occurrence. L'équation 4.35 traduit la densité de probabilité d'attribuer le point  $p$  à la classe  $C_i$ , où  $x_i$  est le nombre de points ayant le label de la classe  $C_i$  parmi les  $K$  plus proches voisins,  $n_i$  est le cardinal de  $C_i$  et  $V^K(p)$  est la plus petite boule contenant les  $K$  plus proches voisins de  $p$ .

$$r_i = \frac{x_i(p)}{n_i \times V^K(p)} \quad (4.35)$$

#### b) L'algorithme de Bayes

L'algorithme de classification de Bayes est un algorithme supervisé paramétrique. Il est basé sur le modèle Bayésien qui repose sur un concept stochastique. La théorie Bayésienne a été introduite par Thomas Bayes en 1764 [Bazeille 2004] et elle consiste, dans la pratique, à modéliser les classes selon les lois gaussiennes. Comme cela est rappelé dans [Meurie 2005], pour chaque pixel donné  $p$ , la classe optimale contenant ce pixel est donnée par la maximisation du critère défini par l'équation 4.36 :

$$f(p, C_i) = -\frac{1}{2}(p - m_i)^T \sum_i^{-1} (p - m_i) - \frac{1}{2} \log |\sum_i| - \log P(C_i) + \frac{k}{2} \log 2\pi \quad (4.36)$$

où  $k$  est le nombre de classes,  $m_i$  est la moyenne des éléments de la classe  $C_i$ ,  $\sum_i$  est la matrice de variance-covariance et  $P(C_i)$  est la probabilité a priori de la classe  $C_i$ . La moyenne et la matrice de covariance sont déterminées à partir des données d'apprentissage.

### c) Les Supports Vecteurs Machines

Les Séparateurs à Vastes Marges (SVM), connu également sous le nom des Supports Vecteurs Machines en anglais, est un classifieur supervisé qui fut introduit par VAPNIK [Vapnik 1996]. Tout comme les deux algorithmes de classification précédents, l'algorithme SVM fait partie de la famille des classifieurs supervisés et requiert donc une phase d'apprentissage permettant d'apprendre les paramètres du modèle de classification approprié pour l'identification des objets d'intérêt dans l'image. Les SVM reposent sur l'utilisation d'une fonction de noyaux (ou kernel en anglais) permettant la séparation optimale (vaste marge) des données. Il existe dans la littérature plusieurs fonctions de noyaux et définitions du critère de similarité [Lafarge 2004, Moussafir 2005]. À titre d'exemple, LAFARGE et al. ont proposé un noyau adaptatif tenant compte des informations de texture et radio-métriques dans le but de reconnaître des zones de feu dans des images SPOT5 de la région PACA [Lafarge 2004]. Néanmoins, la fonction noyau gaussien  $\mathcal{NG}$  (cf. équation 4.37) reste la plus utilisée.

$$\mathcal{NG}(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4.37)$$

## 4.6/ PARAMÉTRAGE DE LA SECONDE APPROCHE PROPOSÉE

Maintenant que nous avons présenté différentes méthodes de simplification d'images et quelques-uns des algorithmes de classification pixellaire (non-supervisés et supervisés) les plus couramment utilisés, nous allons pouvoir tester la deuxième approche proposée dans son ensemble et définir les meilleurs paramètres des algorithmes. Comme indiqué sur la figure 4.5, nous pouvons évaluer, en sortie de l'étape de post-traitement morphologique ou de l'étape de classification pixellaire (pour que le choix des paramètres ne soit pas influencé par cette étape de post-traitement), la qualité de la détection des zones de " ciel " et de " non ciel " de l'image. Pour ce faire, nous avons procédé à une série de combinaisons (méthode de simplification / classifieur) afin de définir les paramètres optimaux et le meilleur couple à utiliser. Dans un premier temps, les tests ont été effectués sur la base d'images du projet LOCOPROL afin d'identifier les deux meilleures méthodes de simplification. Dans un deuxième temps, nous avons ré-itéré une série de tests, mais cette fois sur la base d'images du projet CAPLOC (disposant d'un

grand nombre d'images hétérogènes) afin de définir les deux meilleurs couples (méthode de simplification / classifieur). Tout comme pour la première approche proposée, les résultats ont été évalués avec une méthode d'évaluation adaptée à notre application et utilisant une vérité-terrain.

Comme indiqué précédemment, nous allons dans cette section présenter les résultats de caractérisation de l'environnement de réception des signaux GNSS obtenus avec la seconde approche proposée dans cette thèse. Les tests effectués permettent d'une part, à partir des images de la base du projet LOCOPROL, d'identifier les deux méthodes de simplification d'images les mieux adaptées à notre application (cf. figures 4.26, 4.27, 4.28 et tableau 4.3), et d'autre part, de définir le meilleur couple (simplification/classification), à partir des images de la base du projet CAPLOC, permettant l'identification la plus précise possible des zones " ciel " et " non ciel " dans l'image (cf. figures 4.29, 4.30, 4.31 et tableau 4.4). Il est important de noter que le choix du couple retenu tiendra aussi compte du temps de traitement de l'approche globale (cf. tableau 4.5).

Les figures 4.26, 4.27 et le tableau 4.3 illustrent les différents résultats et le taux de bonne classification (tel que défini dans la section 4.4) obtenus avec les différentes méthodes de simplification (filtres gaussien, médian, moyenneur, RGD, Nagao, exponentiel, Malik-Perona, Outrange et SNN) et classifieurs (KMlocal, FuzzyCmeans, Fisher, SRM, Bayes, Kppv, SVM) précédemment introduits. Nous informons le lecteur que nous avons fait le choix de ne présenter sur ces figures que les différences notables. Par exemple, seulement deux versions du KMlocal sont illustrées ici, sachant que les deux autres versions donnent des résultats quasiment identiques. De la même manière, nous montrons les résultats du classifieur FuzzyCmeans avec  $m = 2$  seulement, les autres valeurs du paramètre de fusion  $m$  n'offrant pas de résultats significativement différents. Cependant, tous les résultats numériques figurent dans le tableau 4.3.

En ce qui concerne les classifieurs non-supervisés (cf. figure 4.26), nous pouvons constater que les résultats obtenus avec KMlocal et FuzzyCmeans sont très proches mais que le filtre médian se départage pour une image (image inrets-4). Pour SRM, les résultats sont en moyenne meilleurs et mettent en avant le filtre Outrange. Pour Fisher, les résultats sont sensiblement identiques, mais le filtre médian offre, là encore, un meilleur taux pour une image donnée (image inrets-3). Il est aussi intéressant de noter que l'utilisation de Fisher sur la composante bleue de l'espace RGB fournit les meilleurs résultats. En ce qui concerne les classifieurs supervisés (cf. figure 4.27), nous pouvons constater que les résultats de classification sont beaucoup plus hétérogènes. Néanmoins, la méthode de simplification qui offre, dans la grande majorité des cas, le meilleur taux de bonne classification, quel que soit le classifieur utilisé, est la RGD. Nous constatons aussi l'importance du choix des paramètres en illustrant sur un exemple que les résultats de classification, pour le classifieur des Kppv, sont sensiblement différents, en terme de pourcentage (parfois près de 6%), selon que nous utilisons un  $nbKppv = 4$  ou un  $nbKppv = 6$ .

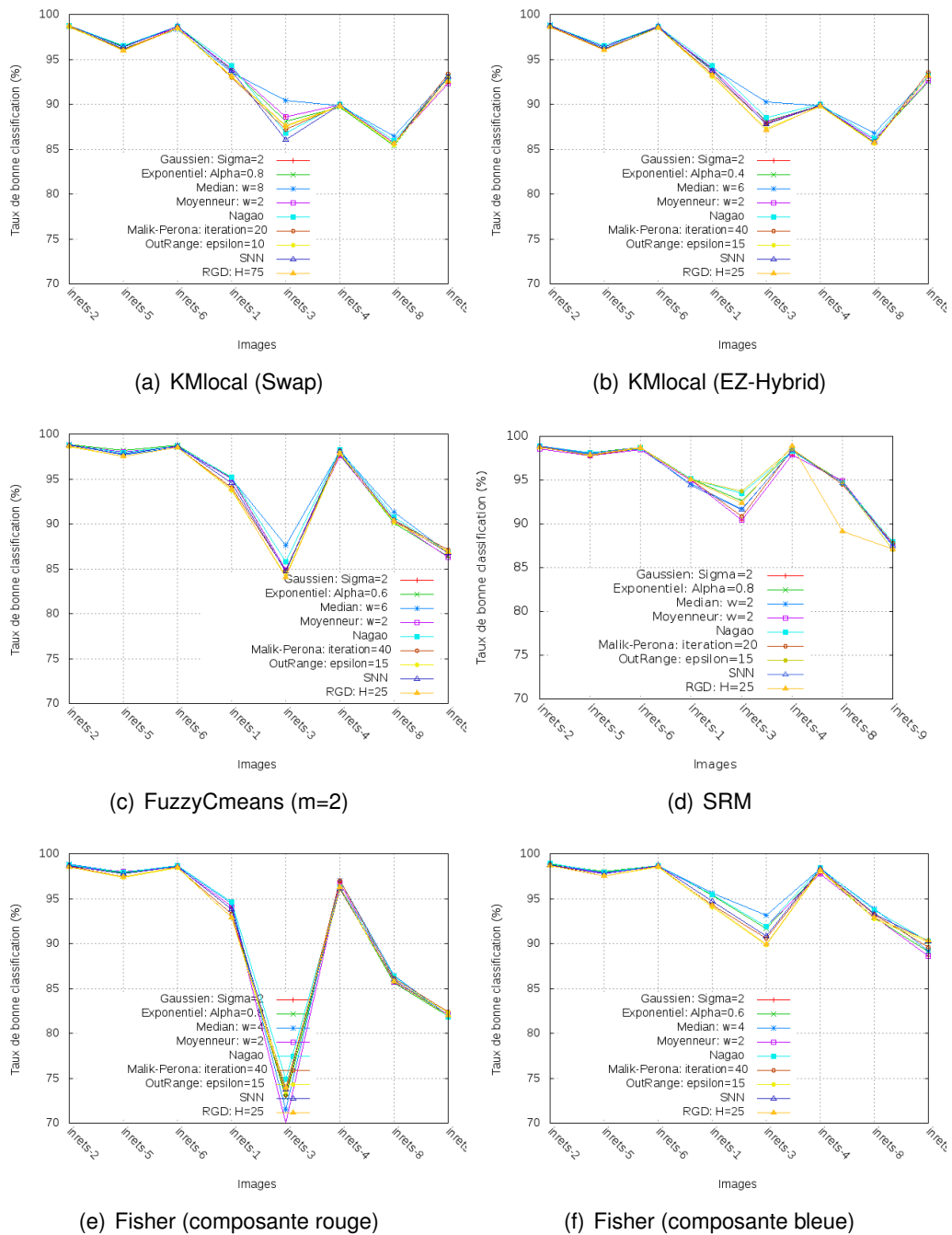


FIGURE 4.26 – Résultats de classification "ciel" / "non-ciel" avec différents algorithmes non-supervisés en fonction de la méthode de simplification utilisée sur la base LOCOPROL.

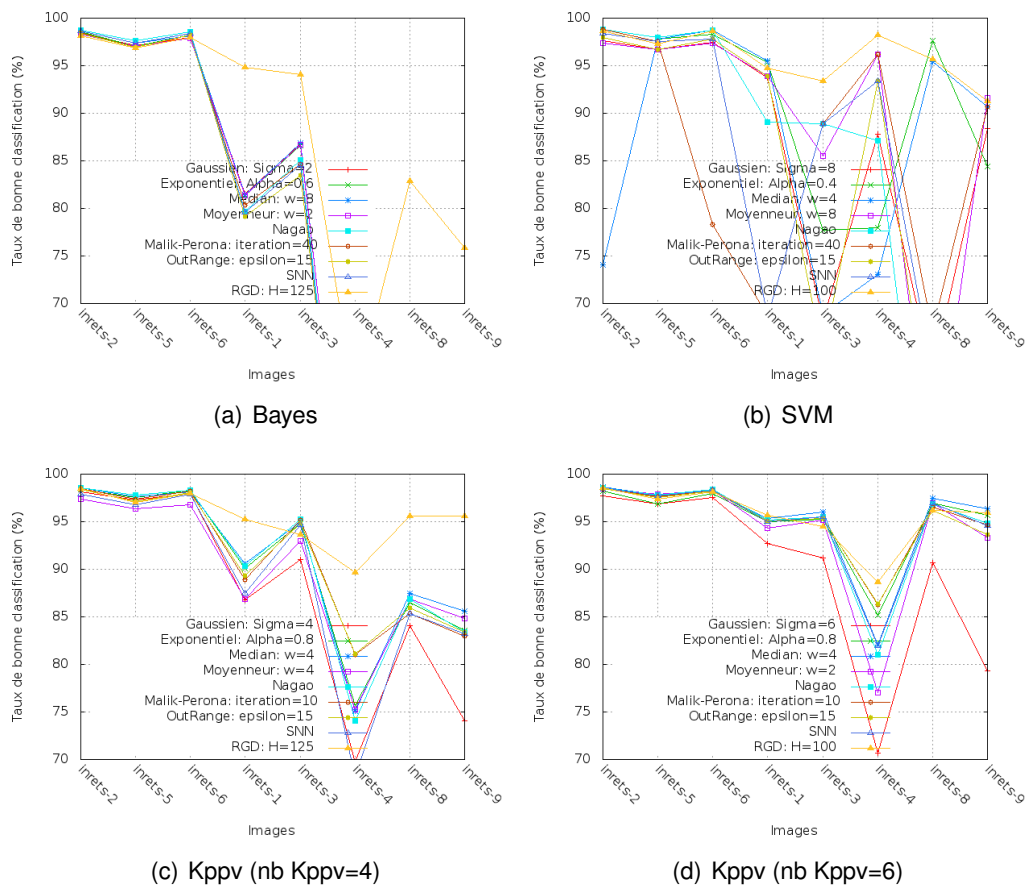


FIGURE 4.27 – Résultats de classification "ciel" / "non-ciel" avec différents algorithmes supervisés en fonction de la méthode de simplification utilisée sur la base LOCOPROL.

Le tableau 4.3 présente le taux de bonne classification moyen "ciel" / "non ciel" sur la base d'images du projet LOCOPROL, en utilisant différents algorithmes de classification (en fonction de différentes valeurs pour les paramètres) et méthodes de simplification (lissage) de l'image. Dans ce tableau, les valeurs en rouge correspondent à la méthode de simplification qui fournit le meilleur taux de bonne classification pour un algorithme de classification donné. En s'appuyant sur les résultats présentés dans ce tableau, nous constatons qu'une simplification effectuée par un filtre médian suivi d'une classification par l'algorithme des KMlocal, FuzzyCmeans ou encore Fisher donnent de très bons résultats. Le meilleur taux de bonne classification (95,7%) est obtenu avec l'algorithme de Fisher sur la composante bleue. Il en est de même pour une simplification faite par une Reconstruction Géodésique par Dilatation (RGD) suivie d'une classification effectuée par l'algorithme des Kppv, Bayes ou SVM. Cette fois-ci, le meilleur taux de bonne classification (96%) est obtenu avec les SVM. Pour une visualisation plus illustrative des résultats, le lecteur pourra se reporter à la figure 4.28 qui présente, pour chaque image, la classification "ciel" (en bleu) / "non ciel" (en marron) obtenue avec la meilleure méthode de simplification pour chaque classifieur considéré. En conclusion, nous retenons le filtre médian et la RGD comme les deux méthodes de simplification les plus intéressantes à utiliser pour la deuxième série de tests effectuée sur la base d'images du projet CAPLOC, l'objectif de cette deuxième série de tests étant de déterminer le meilleur couple (simplification/classifieur).



		Filtres															
		Gaussien		Median		Moyenneur		RGD		Nagao	Exponentiel		Malik -Perona		OutOfRange		SNN
		P	$\sigma$	P	$W$	P	$W$	P	H	P	P	$\alpha$	P	$itr$	P	Seuil	P
Fisher	Rouge	90.5	2	90.8	4	90.5	2	90.7	25	91.2	90.7	0.8	90.8	40	85.3	150	90.8
	Vert	94.4	2	94.9	4	94.4	2	94.1	50	94.7	94.5	0.6	94.3	40	94.4	200	94.3
	Bleue	95.2	2	95.7	4	95.2	2	95	25	95.7	95.3	0.6	95.1	40	95.3	200	95.3
	Moyenne (R,G,B)	93.6	2	94.2	6	93.6	2	93.3	25	93.9	93.6	0.6	93.5	40	93.6	200	93.6
KMlocal	EZHybrid	93.1	2	93.4	6	93.1	2	92.8	25	93.3	93	0.4	93	40	93	150	93
	Hybrid	93.1	2	93.5	6	93.1	2	92.8	25	93.3	93	0.6	93	40	92.9	150	93
	Loyds	93.1	2	93.4	6	93.1	2	92.8	25	93.3	93	0.4	93	40	93	150	93
	Swap	93.1	2	93.3	8	93.1	2	92.8	75	93	93	0.8	92.8	20	83.7	150	92.8
Fuzzy Cmeans	m = 1	93.7	2	94.3	6	93.7	2	93.5	25	94	93.8	0.6	93.6	40	93.7	200	93.7
	m = 1.25	93.7	2	94.3	6	93.7	2	93.5	25	94.1	93.8	0.6	93.6	40	93.7	200	93.7
	m1.5	93.7	2	94.3	6	93.7	2	93.5	25	94.1	93.8	0.6	93.6	40	93.7	200	93.73
	m = 1.75	93.7	2	94.3	6	93.7	2	93.5	25	94.1	93.8	0.6	93.7	40	93.7	200	93.7
	m = 2	93.7	2	94.2	6	93.7	2	93.5	25	94	93.7	0.6	93.7	40	93.7	200	93.7
SRM		95.1	2	95.3	2	95.1	2	94.8	25	95.7	95.5	0.8	95.2	20	95.6	200	95.3
Bayes		76.9	2	76.9	8	76.8	2	87.6	125	76	77	0.6	76.1	40	76.8	200	75.9
Kppv	nbKppv4	87.4	4	91	4	89.7	4	95.4	125	90.6	90.6	0.8	90.9	10	90.5	50	88.9
	nbKppv6	93.1	2	95.2	4	93.9	2	95.8	50	94.8	95.1	0.8	95.3	10	94.6	50	94.8
SVM		86.6	8	86.8	4	88.7	8	96	100	81.4	87.9	0.4	85.8	40	85.9	150	84.6

TABLE 4.3 – Résultats de classification " ciel " / " non ciel " avec les différents algorithmes de classification, méthodes de simplification utilisés et paramètres, sur la base d'images du projet LOCOPROL.

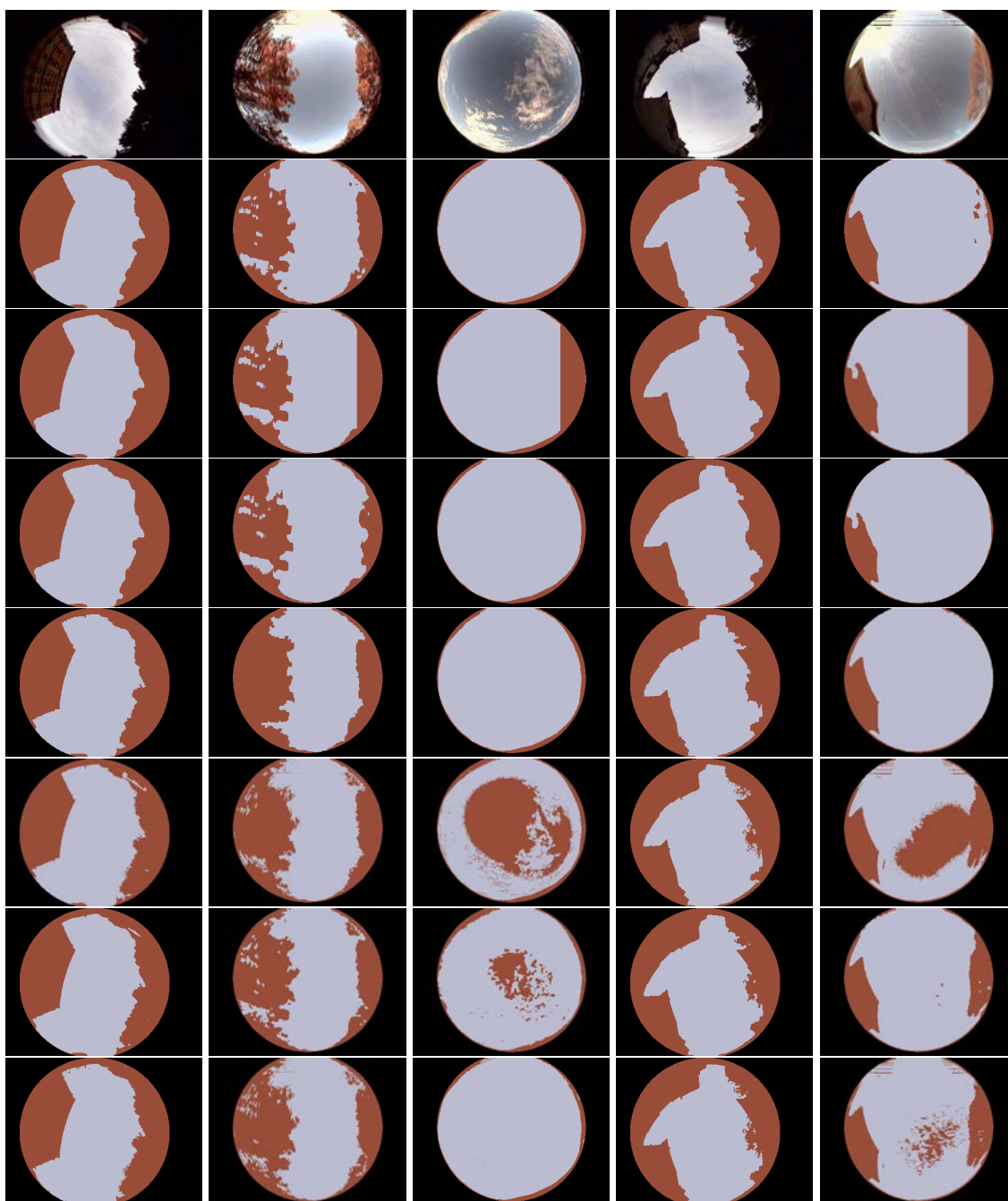
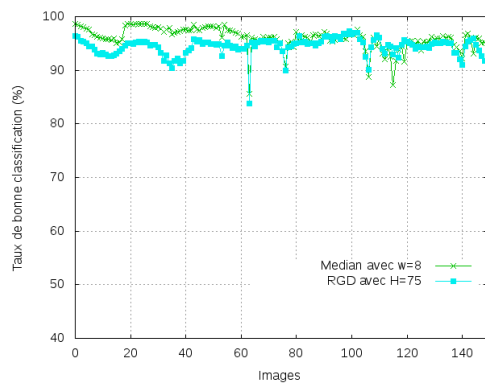


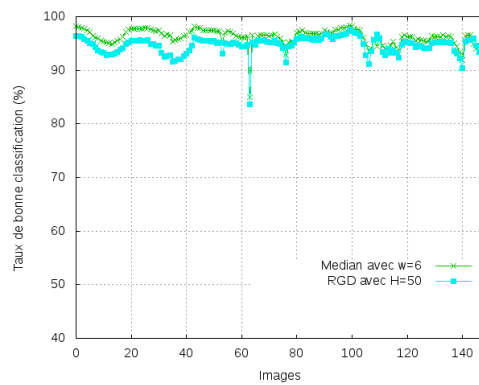
FIGURE 4.28 – Résultats de classification " ciel " (en bleu) / " non ciel " (en marron) sur 5 images de la base d'images du projet LOCOPROL avec différents classifieurs (de haut en bas : Fisher (composante couleur bleue), KMlocal (version Hybrid), FuzzyCmeans (m=2), SRM, Bayes, Kppv (nbKppv=6), SVM).

Les premiers tests ayant permis l'identification des deux méthodes de simplification les plus adaptées à notre application, nous avons ensuite cherché à définir le meilleur couple simplification/classification sur la base d'images du projet CAPLOC. Les figures 4.29 et 4.30 illustrent le taux de bonne classification obtenu en fonction du classifieur utilisé. Pour les classifieurs non-supervisés, nous remarquons que les résultats sont sensiblement identiques, que nous utilisons un filtre médian ou une RGD, hormis pour Fisher (sur la

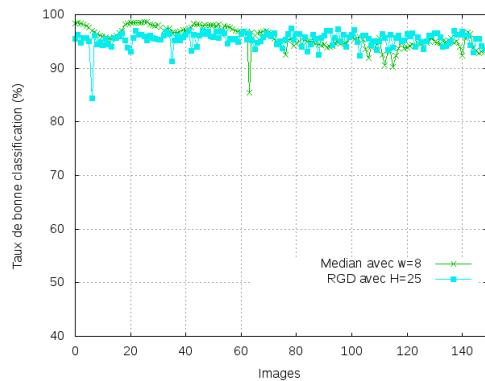
composante rouge) et SRM qui montrent quant à eux une très nette amélioration avec l'utilisation de la RGD (jusqu'à 24%). Pour les classifieurs supervisés, les résultats sont eux aussi très proches (hormis pour les SVM), mais montrent que la RGD offre dans tous les cas des résultats supérieurs à ceux obtenus avec un filtre médian.



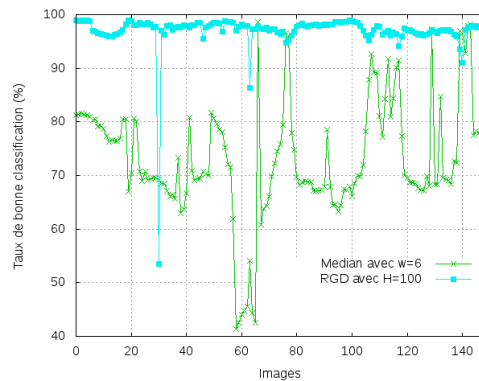
(a) KMlocal (Swap)



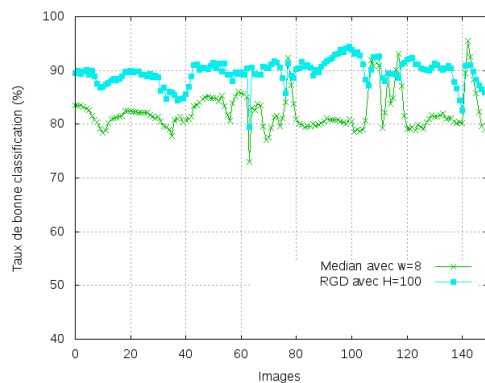
(b) KMlocal (EZ-Hybrid)



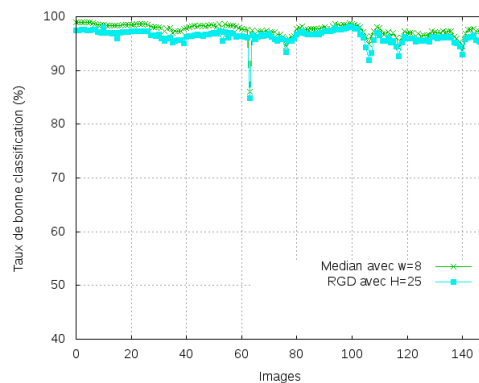
(c) FuzzyCmeans (m=2)



(d) SRM



(e) Fisher (composante rouge)



(f) Fisher (composante bleue)

FIGURE 4.29 – Résultats de classification ciel/non-ciel avec différents algorithmes non-supervisés en fonction de la méthode de simplification utilisée sur la base CAPLOC.

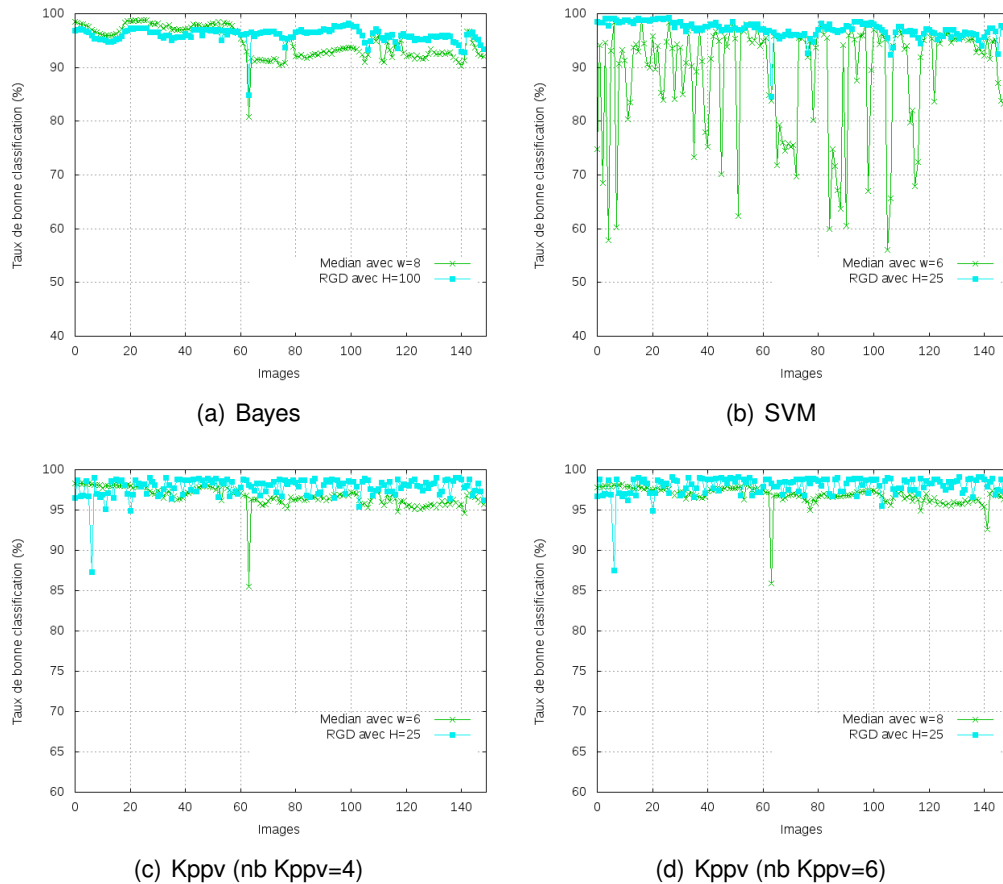


FIGURE 4.30 – Résultats de classification ciel/non-ciel avec différents algorithmes supervisés en fonction de la méthode de simplification utilisée sur la base CAPLOC.

Le tableau 4.4 présente le taux de bonne classification moyen " ciel " / " non ciel " sur la base d'images du projet CAPLOC, en utilisant différents algorithmes de classification (en fonction de différentes valeurs pour les paramètres) et les deux méthodes de simplification (retenues aux premiers tests : filtre médian et RGD). Ce tableau permet également d'illustrer l'apport de l'étape de simplification et l'étape de post-traitement dans la stratégie globale proposée. En s'appuyant sur les résultats présentés dans le tableau 4.4, nous pouvons constater, que dans un cadre non supervisé, une classification réalisée par l'algorithme de Fisher (sur la composante bleue de l'espace RGB) donne les meilleurs résultats (i.e. 97,5% avant post-traitement et 98,1% après post-traitement) si elle est couplée à une simplification par un filtre médian, et arrive en deuxième position si elle est associée à une simplification par RGD (i.e. 96,4% avant post-traitement et 97,2% après post-traitement). Dans un cadre supervisé, l'utilisation d'une RGD offre toujours les meilleurs résultats avant le post-traitement, et prend la première position dans 3 cas sur 4 cas par rapport à l'utilisation d'un filtre médian avec post-traitement. L'utilisation de la RGD couplée à l'algorithme des Kppv (avec un nbKppv = 6) permet d'obtenir le meilleur taux de bonne classification, à savoir 98.1%. Là encore, pour une visualisation plus illustrative, la figure 4.31 présente sur cinq images de la base CAPLOC, la classification " ciel " (en bleu) / " non ciel " (en marron) obtenue avec la meilleure méthode de simplification pour chaque classifieur utilisé.

	Sans simplification	Sans post-traitement				Avec post-traitement			
		Median		RGD		Median		RGD	
	P (%)	P (%)	W	P (%)	H	P (%)	W	P (%)	H
	Classification non-supervisée								
	Fisher								
Rouge	81.3	82.3	8	89.6	100	86.9	8	91.6	100
Vert	89.9	91.7	8	93.5	50	92.9	8	93.9	50
Bleue	95.8	97.5	8	96.4	25	98.1	8	97.2	25
Moyenne (R,G,B)	91.4	92.8	8	93.6	50	93.3	8	94.1	50
KMlocal									
EZHybrid	94.7	96.4	8	95.2	25	96.5	8	96.4	25
Hybrid	94.7	96.4	8	95.2	25	96.5	8	96.3	25
Loyds	94.7	96.4	8	95.2	25	96.5	8	96.5	25
Swap	94.3	96.1	8	95.4	25	95.7	8	96	25
Fuzzy Cmeans									
m=1	95.2	96.4	8	95.3	25	97.7	8	96.2	25
m=1.25	95.1	96.4	8	95.3	25	97.6	8	97.3	25
m=1.5	94.4	96.2	8	95.4	25	97.5	8	97.5	25
m=1.75	94.8	95.9	8	95.4	25	97.2	8	97.9	25
m=2	94.6	95.7	8	95.4	25	96.9	8	97.9	25
SRM									
	75.7	73.18	6	97	100	97	6	97	100
Classification supervisée									
Bayes									
	90.3	94.6	8	96	100	97.5	8	97.3	100
Kppv									
nbKppv=4	97.4	96.7	6	97.8	25	97	6	98	25
nbKppv=6	97.4	96.8	8	98	25	97.5	8	98.1	25
SVM									
	94.7	88.8	6	97.1	25	88.8	6	97.1	25

TABLE 4.4 – Résultats de classification " ciel " / " non ciel " avec différents algorithmes de classification, méthodes de simplification utilisés et paramètres, sur la base d'images du projet CAPLOC.

Le tableau 4.5 illustre quant à lui, le temps de traitement (avec un PC Core TM2 cadencé 2.80 Hz) de l'approche adoptée en fonction du couple (méthode de simplification / algorithme de classification) utilisé, appliquée sur des images de taille 640 × 512 pixels.

	Fisher	KMlocal	Fuzzy Cmeans	SRM	Bayes	Kppv	SVM
Median (en s)	0.035	0.41	3.30	0.35	0.34	15.3	5.0
RGD (en s)	0.037	0.45	3.34	0.47	0.47	15.4	5.04

TABLE 4.5 – Temps de traitement de l'approche en fonction du couple méthode de simplification / algorithme de classification.

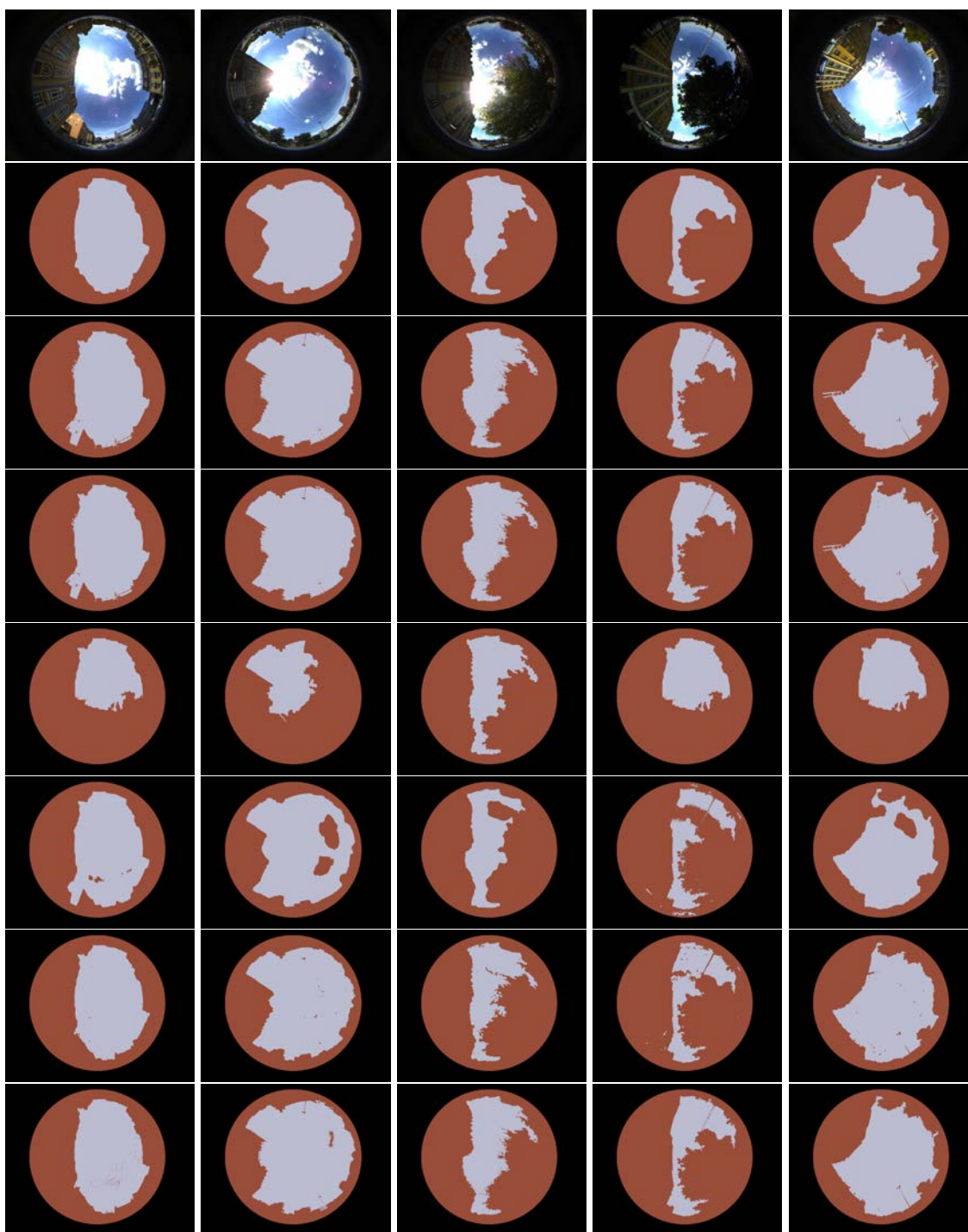


FIGURE 4.31 – Résultats de classification " ciel " (en bleu) / " non ciel " (en marron) sur 5 images de la base d'images du projet CAPLOC avec différents algorithmes de classification (de haut en bas : Fisher (composante couleur bleue), KMLocal (version Loyds), FuzzyCmeans (m=2), SRM, Bayes, Kppv (nbKppv=6), SVM).

Nous remarquons que le temps de traitement de l'approche proposée est peu affecté par la méthode de simplification utilisée, mais peut être jusqu'à 400 fois supérieur selon le classifieur employé (par exemple entre Fisher et Kppv). L'application du projet CAPLOC



nous imposant de considérer la contrainte temps-réel, nous retiendrons finalement l'algorithme de Fisher (avec la composante bleue) comme meilleur classifieur. Une première préférence est donnée au filtre médian par rapport à la RGD, mais ce choix doit être confirmé sur les images (qui seront plus nombreuses) de la prochaine campagne d'acquisition du projet CAPLOC prévue en fin d'année 2013. Le temps de traitement de la stratégie globale s'élève 35 ms/image.

## 4.7/ CONCLUSION

Dans ce chapitre, nous avons proposé et développé deux approches basées sur le traitement d'image, permettant de caractériser l'environnement de réception des signaux GNSS.

La première approche est basée sur une segmentation de l'image par combinaison des informations de couleur et de texture. Cette méthode de segmentation, détaillée dans le chapitre 3, se place parmi les meilleurs méthodes de segmentation de la littérature. Elle est ensuite couplée à une classification de régions en deux classes afin d'obtenir les zones " ciel " / " non ciel " de l'image. Les résultats obtenus sont très satisfaisants, avec un taux de bonne classification de 95.7% sur la base d'images du projet LOCOPROL et 97.9% sur la base d'images du projet CAPLOC. Cependant, cette première approche présente l'inconvénient de nécessiter un coût calculatoire incompatible avec la contrainte temps-réel imposée par notre application.

La deuxième approche proposée consiste tout d'abord à simplifier l'image initiale par un filtre adapté afin de lisser celle-ci et atténuer certains artefacts dus au système de prise de vues (caméra avec objectif fisheye) et les pics qui apparaissent suite aux variations brusques de la luminosité. Cette étape de simplification est suivie d'une étape de classification qui, contrairement à la première approche proposée, traite directement les pixels (et non pas les régions issues de la segmentation) pour lui attribuer la classe " ciel " ou " non ciel ". Enfin une étape de post-traitement, basée sur des opérations de morphologie mathématique, permet de reclasser les pixels erronés. Cette deuxième approche offre un taux de bonne classification qui atteint 98.1%. Elle permet par ailleurs une utilisation en temps-réel par rapport à notre application.

Les résultats de détection " ciel " / " non ciel " étant quasiment identiques, nous avons retenu pour la suite la deuxième approche qui présente un coût calculatoire compatible avec la contrainte temps-réel imposée par notre application.

L'objectif de la caractérisation de l'environnement de réception des signaux GNSS consiste à pouvoir déterminer, après repositionnement, les satellites se trouvant dans la région zone " ciel " de l'image (ie. les satellites dont les signaux GNSS sont reçus via des trajets directs). Ceci permettra d'exclure les satellites situés dans une zone " non ciel " (ie. ceux dont les signaux GNSS sont réfléchis ou bloqués) lors du calcul de la nouvelle position (se voulant plus précise). Ce travail fera l'objet du chapitre 5.

# UTILISER L'IMAGE POUR LOCALISER AVEC PLUS DE PRÉCISION

## Sommaire

<b>5.1</b>	<b>Introduction</b>	<b>181</b>
<b>5.2</b>	<b>Introduction à la localisation par satellites</b>	<b>182</b>
5.2.1	Le système GNSS	183
5.2.2	Les données de mesure	183
5.2.3	Le calcul de la position	183
<b>5.3</b>	<b>Impact de l'environnement de réception sur la précision de la localisation</b>	<b>184</b>
5.3.1	Définition du DOP	184
5.3.2	L'imprécision de la mesure de pseudo-distance causée par les multi-trajets	185
5.3.3	Méthodes de mitigation du problème de multi-trajet	186
<b>5.4</b>	<b>Application de techniques de traitement d'images pour la localisation</b>	<b>187</b>
5.4.1	Repositionnement des satellites dans l'image	188
5.4.2	Détection de la ligne d'horizon	190
<b>5.5</b>	<b>Calcul d'une nouvelle position</b>	<b>191</b>
5.5.1	Systèmes d'acquisition	191
5.5.2	Description des données expérimentales	191
5.5.3	Calcul de position à partir des satellites LOS uniquement	192
5.5.4	Utilisation combinée du DOP et de l'exclusion des satellites NLOS	196
5.5.5	Discussions sur l'impact de la végétation sur la visibilité satellitaire	198
5.5.6	Analyse quantitative	200
<b>5.6</b>	<b>Conclusion</b>	<b>204</b>

## 5.1/ INTRODUCTION

Ce dernier chapitre est un chapitre applicatif. Nous y présentons comment une nouvelle stratégie basée sur l'analyse d'images permet d'estimer la position d'un véhicule en mouvement de manière précise. Les signaux satellitaires subissent des phénomènes de masquage et une propagation perturbée par les obstacles proches de l'antenne en environnement urbain. Le calcul de la position reposant sur la mesure du temps de propagation, ces phénomènes dégradent la précision de la localisation. L'approche suivie dans ce chapitre consiste à éliminer du calcul de position les signaux qui dégraderaient cette information pour ne garder que les signaux utiles. L'objectif est de montrer l'intérêt

de ce choix et de discuter ses limites. On distingue pour cela deux types de signaux reçus : les signaux qui peuvent être reçus via un trajet direct (LOS - Line-Of-Sight) et les signaux dont le trajet est bloqué, ou qui peuvent être reçus avec une atténuation trop forte (reçus après réflexions sur des obstacles : phénomène de multi-trajet) pour être exploité (NLOS - Non-line-Of-Sight).

Dans ce chapitre, nous allons comparer trois stratégies d'estimation de la position à partir d'un même jeu de satellites reçus. La première stratégie sert de référence de fonctionnement. Elle consiste à appliquer un calcul de position " classique " où il s'agit d'estimer une position en tenant compte des signaux reçus par tous les satellites de la constellation disponible. Partant de l'idée d'exclure toutes les sources de multi-trajet, la deuxième stratégie consiste à estimer une position en utilisant uniquement les signaux reçus en LOS (ne subissent pas de réflexions par les obstacles de l'environnement de réception de signaux GNSS). La détection de ces signaux n'est cependant pas facile, dans la mesure où cette information n'est pas observable. Ces derniers peuvent être identifiés par un seuillage du rapport signal sur bruit en choisissant un seuil empirique. Dans les travaux de VIANDIER, on utilise le CN0 (Carrier-to-Noise ratio) avec un seuil égal à 35dB [Viandier 2011]. Dans le cadre du projet CAPLOC (Projet PREDIT financé par le MEDDE/DRI), nous proposons d'identifier les satellites LOS et NLOS, à chaque instant, à l'aide de l'analyse d'une séquence d'images acquise simultanément avec les données GPS. Suite à la phase de traitement d'images (cf. chapitre 4), nous disposons d'une image constituée de deux classes " ciel " et " non ciel ", ce qui nous permet de classer les satellites en LOS s'ils se trouvent dans les régions " ciel " et en NLOS sinon. Nous avons testé les performances obtenues lorsque les satellites NLOS, dont les mesures sont potentiellement dégradées, sont exclus. Ces résultats, comparés aux mesures utilisant tous les satellites disponibles, nous ont conduit à proposer une troisième stratégie d'estimation de position qui consiste à utiliser l'information DOP (Dilution of Precision), caractérisant la géométrie des satellites. L'information DOP est donc exploitée afin d'appliquer la stratégie d'exclusion seulement lorsqu'elle est appropriée.

Ce chapitre est organisé de la manière suivante : Dans un premier temps, nous introduisons les principes de fonctionnement d'un système GNSS et de calcul d'une position. Nous présentons ensuite comment l'usage de l'image permet d'améliorer la précision de la localisation dans les zones de réception contraintes. Après projection des satellites dans l'image, nous détectons via un traitement d'images, basé sur la morphologie mathématique, la ligne d'horizon afin de déterminer les satellites LOS et les satellites NLOS. Cette identification étant achevée, nous comparons différentes techniques de calcul de position destinées à optimiser l'usage des satellites disponibles pour fournir une position la plus précise possible.

## 5.2/ INTRODUCTION À LA LOCALISATION PAR SATELLITES

Les systèmes de localisation par satellites (GNSS - Global Satellite Navigation System) reposent sur le principe de la mesure du temps de propagation des signaux émis par les satellites de la constellation jusqu'à la réception. Cette section a pour objet d'introduire les principes fondamentaux permettant par la suite de comprendre la démarche proposée

pour l'amélioration de la précision lors du calcul d'une position d'un mobile.

### 5.2.1/ LE SYSTÈME GNSS

Un système GNSS est un système composé d'un ensemble de satellites et d'un système sol de contrôle et maintenance. Le système GNSS le plus connu est le GPS (Global Positioning System) entretenu par les Etats-Unis. Il en existe d'autres systèmes tels que GLONASS (le système russe) et bientôt GALILEO (le système européen) et BEIDOU (le système chinois). Les satellites de localisation sont des satellites MEO (Medium Earth Orbit) défilants autour de la terre. Leur nombre (une trentaine environ) permet d'assurer une couverture mondiale.

### 5.2.2/ LES DONNÉES DE MESURE

Les satellites émettent en continu des signaux que le récepteur utilise pour extraire leur temps de parcours et un certain nombre d'informations utiles contenues dans le message de navigation. Ce temps de parcours, multiplié par la vitesse de propagation des ondes  $c$ , permet d'estimer la distance entre chacun des satellites et le récepteur. On appelle cette mesure la pseudo-distance. Cette mesure  $\rho$  est entachée d'un certain nombre d'erreurs que l'on retrouve dans l'équation 5.1.

$$\rho = d + c\delta t + I + T + m + \epsilon \quad (5.1)$$

où  $d$  est la distance géométrique,  $\delta t$  le décalage en temps entre l'horloge satellite et l'horloge récepteur.  $I$  et  $T$  sont respectivement les erreurs liées à la propagation des signaux dans l'ionosphère et la troposphère. Elles peuvent être corrigées par différentes techniques (utilisation de récepteurs bi-fréquences ou de modèles mathématiques) que nous ne développerons pas ici.  $\epsilon$  est le bruit du récepteur.  $m$  est le bruit causé par les multi-trajet et  $\epsilon$  le bruit du récepteur. Il est directement lié à l'environnement de réception et peut difficilement être modélisé par une loi générale. C'est ce poste d'erreur que nous traitons dans ce chapitre.

### 5.2.3/ LE CALCUL DE LA POSITION

La position du récepteur est calculée par le récepteur lui-même à partir de la réception d'au moins 4 signaux simultanément, en utilisant le principe dit trilatération. Mathématiquement, il s'agit de résoudre un système d'équations à 4 inconnues que sont la position du récepteur  $\{x, y, z\}$  et le décalage d'horloge  $\delta t$ . Les méthodes les plus classiques sont la méthode des moindres carrés et le filtre de Kalman Etendu. La méthode des moindres carrés fournit la solution la plus simple, mais ne prend pas en compte les informations disponibles sur la mesure pour corriger l'estimation de la position (ce que fait le filtre de Kalman). Le filtre de Kalman Etendu que nous utilisons dans ce chapitre est celui implémenté par VIANDIER [Viandier 2011]. Ce dernier a constaté que le filtre de Kalman classique n'est pas exploitable, étant donné que l'expression de la pseudodistance est faiblement non-linéaire. Pour cette raison, il a eu recours à des approximations permettant la linéarisation des fonctions d'état et d'observation dans l'expression du filtre de Kalman classique.

### 5.3/ IMPACT DE L'ENVIRONNEMENT DE RÉCEPTION SUR LA PRÉCISION DE LA LOCALISATION

La précision de la localisation obtenue est directement liée à la qualité des mesures réalisées. Deux facteurs liés à l'environnement sont à mentionner particulièrement :

- la géométrie des satellites utilisés, caractérisée par le DOP (Dilution of Precision).
- l'imprécision de la mesure de pseudo-distance causée par les multi-trajets.

#### 5.3.1/ DÉFINITION DU DOP

Le DOP (Dilution of Precision) constitue un indicateur permettant de donner une idée sur la répartition géométrique des satellites à un instant donné. Lorsque le récepteur reçoit les signaux émis par les satellites, sa position peut être déterminée par l'intersection des sphères dont les centres sont les positions des satellites. Le DOP est généralement proportionnel à l'inverse du volume du polyèdre formé par le récepteur et les satellites dont les signaux sont reçus par ce récepteur. Lorsque les satellites sont éloignés les uns des autres (cf. figure 5.1(a)), la précision de la position estimée est meilleure que lorsque ceux-ci sont très regroupés dans un même volume au dessus de l'antenne (cf. figure 5.1(b)). Par conséquent, la valeur du DOP est un indicateur de la précision du positionnement. Plus le DOP est faible, meilleure est la précision.

Le DOP est calculé à partir de la matrice de covariance de la position donnée dans un repère plan ou ENU (East North Up). Il est important de noter qu'il prend en données d'entrée, les positions des satellites dans un référentiel géocentré. Il existe plusieurs types de DOP :

- **GDOP** (Global DOP)

$$GDOP = \frac{\sqrt{\sigma_e^2 + \sigma_n^2 + \sigma_u^2 + \sigma_t^2}}{\sigma} \quad (5.2)$$

- **PDOP** (Precision DOP)

$$PDOP = \frac{\sqrt{\sigma_e^2 + \sigma_n^2 + \sigma_u^2}}{\sigma} \quad (5.3)$$

- **HDOP** (Horizontal DOP)

$$HDOP = \frac{\sqrt{\sigma_e^2 + \sigma_n^2}}{\sigma} \quad (5.4)$$

- **VDOP** (Vertical DOP)

$$VDOP = \frac{\sigma_u}{\sigma} \quad (5.5)$$

- **TDOP** (Time DOP)

$$TDOP = \frac{\sigma_t}{\sigma} \quad (5.6)$$

où  $\sigma_e^2$ ,  $\sigma_n^2$  et  $\sigma_u^2$  sont les variances de la position dans le repère ENU et  $\sigma_t^2$  désigne la variance du temps. Le calcul du DOP est basé sur l'hypothèse que les satellites sont en état LOS. La technique de calcul de ce critère est introduite avec plus de détails dans les travaux de VIANDIER [Viandier 2011]. Pour évaluer la qualité du PDOP, 4 seuils peuvent être définis :

- si PDOP est inférieur à 3, l'incertitude sur la position du récepteur est faible.
- si PDOP est inférieur à 6, l'incertitude sur la position du récepteur est moyenne.

- si PDOP est inférieur à 9, l'incertitude sur la position du récepteur est large.
- si PDOP est supérieur à 9, la position estimée est à exclure.

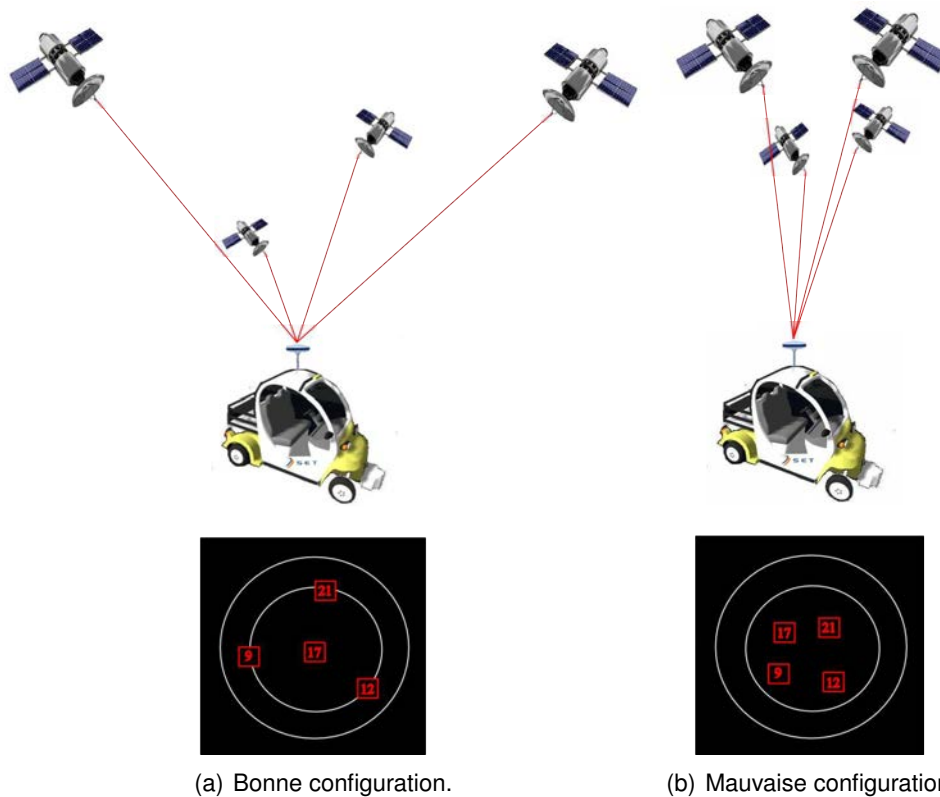


FIGURE 5.1 – Exemple d'une bonne et d'une mauvaise configuration géométrique des satellites.

### 5.3.2/ L'IMPRÉCISION DE LA MESURE DE PSEUDO-DISTANCE CAUSÉE PAR LES MULTI-TRAJETS

En plus des perturbations classiques mentionnées plus haut, l'environnement proche de l'antenne peut contraindre fortement la réception des signaux et la qualité de la mesure de la pseudo-distance. Le premier impact de cet environnement peut être le blocage des signaux émis. Lorsque les signaux sont complètement stoppés par les obstacles environnants, le récepteur n'est pas en mesure de les utiliser. Par conséquent, le nombre de signaux disponibles diminue, et lorsqu'il devient inférieur à 4, le calcul de la position n'est alors plus possible. On parle alors d'indisponibilité du service.

Le signal peut également être reçu en plusieurs échos, d'amplitudes et de retards différents. La mesure du temps de propagation du signal est alors plus délicate. La figure 5.2 illustre un exemple de réception d'un signal en NLOS après une ou plusieurs réflexions. La mesure de temps doit ainsi inclure le retard induit par le trajet supplémentaire du signal réfléchi.



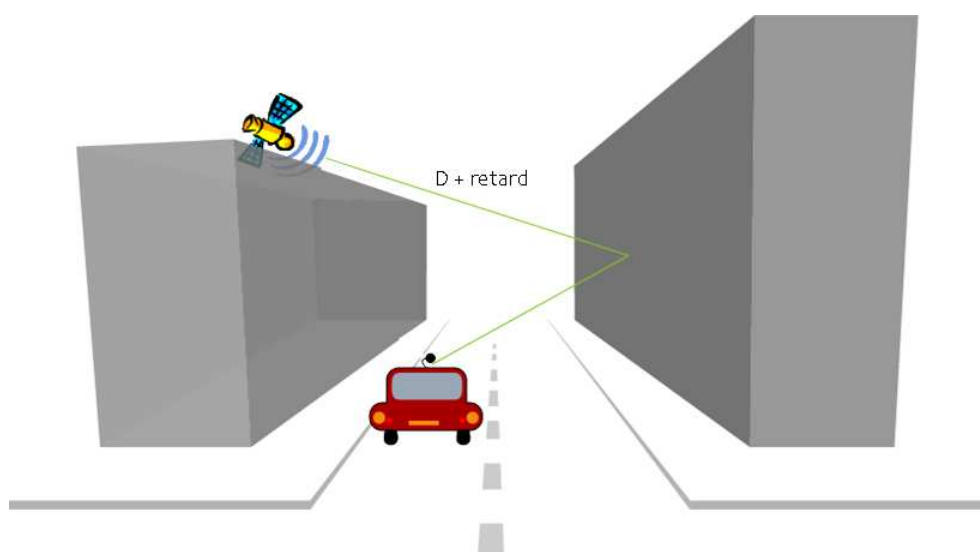


FIGURE 5.2 – Signal reçu à partir d'un satellite NLOS, après une réflexion.

### 5.3.3/ MÉTHODES DE MITIGATION DU PROBLÈME DE MULTI-TRAJET

Cette thèse ne porte pas spécifiquement sur ce thème. Nous ne présentons donc pas ici un inventaire exhaustif des techniques de mitigation disponibles dans la littérature. Il existe en effet une multitude de points d'entrée pour ce travail tout au long de la chaîne de réception et de traitement des signaux GNSS [Lentmaier 2008] : traitement d'antennes, modélisation du canal de propagation, développement de nouveaux corrélateurs, algorithmes de navigation prenant en compte les bruits induits, ajouts de sources supplémentaires tels que les capteurs inertiels ou même les cartographies numériques. Dans cette thèse et dans le cadre du projet CAPLOC, le choix de la méthode s'est porté sur la détection et l'exclusion des satellites potentiellement perturbateurs de la qualité de la mesure. En environnement contraint, tel qu'un environnement urbain, les principaux perturbateurs sont les obstacles proches de l'antenne de réception (le terme " proche " étant relatif face aux 20000 km d'altitude des satellites).

Ces obstacles ont un impact direct sur l'amplitude des signaux reçus, répercutée sur la valeur du rapport signal sur bruit (CN0). Dans les travaux de VIANDIER [Viandier 2011] et MARAIS [Marais 2012], une solution, basée sur le seuillage de l'amplitude du signal GNSS reçu, est proposée. La difficulté de la détection des satellites LOS/NLOS sur le critère de seuillage réside dans le choix de la valeur du seuil CN0. Ce dernier peut être sensiblement différent d'un équipement à l'autre. De plus, il est fixé à partir de données statistiques et la dispersion des valeurs autour de ce seuil implique de mauvaises détections. C'est pourquoi d'autres techniques sont aujourd'hui étudiées. En effet, l'IFSTTAR-LEOST a développé, dans cet objectif, l'outil PREDISSAT [Marais 2002], fondé sur l'utilisation de caméras classiques et d'un processus de post-traitement. MERGURO [Meguro 2008] a repris le concept avec une caméra Infra-Rouge. Avec le développement des modèles 3D urbains, d'autres auteurs [Groves 2011, Bourdeau 2012, Obst 2012] ou encore [Peyret 2011], ont publié des résultats montrant comment ces modèles pouvaient également permettre de déterminer la présence ou non de trajets directs, en comparant la position des obstacles modélisés et la position des satellites. BOURDEAU

[Bourdeau 2012] estime le biais affecté aux pseudo-distances et causé par la réflexion des signaux NLOS, afin d'en tenir compte dans le processus d'estimation de la position. Dans notre approche, nous adoptons, comme dans [Obst 2012], une stratégie d'exclusion des satellites NLOS. L'approche proposée fera l'objet de la section suivante.

## 5.4/ APPLICATION DE TECHNIQUES DE TRAITEMENT D'IMAGES POUR LA LOCALISATION

À l'issue du travail présenté au chapitre 4, nous disposons d'une caractérisation de l'environnement de réception des signaux GNSS (une classification de l'image en une zone " ciel " et une zone " non ciel "). Nous allons, dans la suite de ce chapitre, détailler comment utiliser cette information pour tenter d'améliorer la précision de la position d'un mobile en milieu urbain

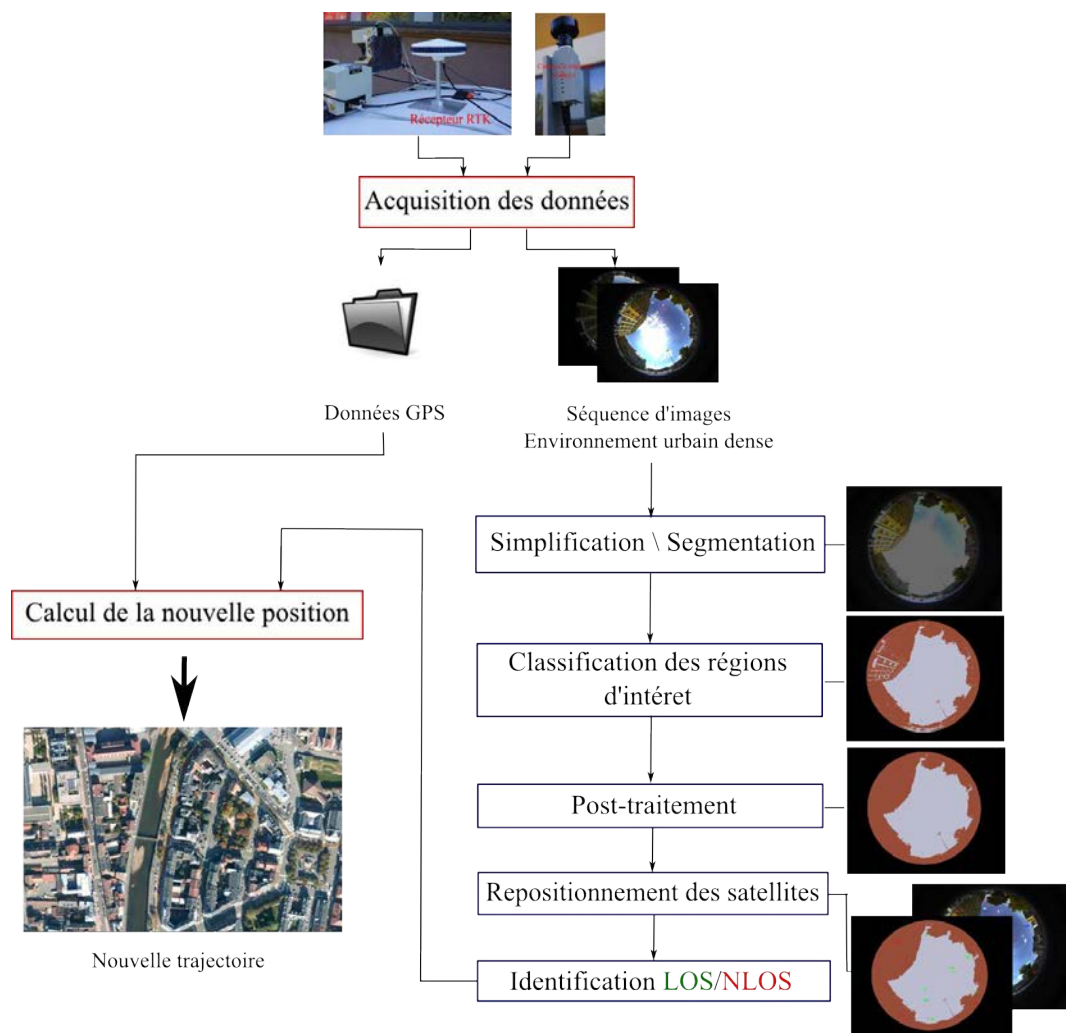


FIGURE 5.3 – Synopsis de la stratégie globale pour l'amélioration de la précision lors du calcul de la position d'un récepteur GNSS.

Le synopsis illustré sur la figure 5.3 décrit la stratégie globale proposée. Elle est

composée d'une brique d'acquisition de données, de cinq briques de traitement d'images et d'une dernière brique de localisation (calcul de position). La première brique concerne la phase d'acquisition des données de localisation (à l'aide des récepteurs GNSS embarqués) et des données images (à l'aide d'une caméra couleur à objectif fisheye installée sur le toit du véhicule et orientée verticalement vers le ciel). Les trois premières briques de traitement d'images (simplification/segmentation, classification pixellaire et post-traitement) correspondent à la phase de caractérisation de l'environnement de réception des signaux GNSS, développée dans le chapitre 4. Les deux dernières briques de traitement d'images concernent respectivement le repositionnement des satellites dans l'image et l'identification des satellites LOS/NLOS (ie. se situant dans une zone de " ciel " / " non ciel "). Ces deux briques seront détaillées dans ce chapitre. La dernière brique de l'approche proposée correspond au calcul d'une nouvelle position de localisation, en excluant/pondérant les différents satellites suivant leur position dans l'image. Cette dernière phase sera détaillée dans ce chapitre.

#### 5.4.1/ REPOSITIONNEMENT DES SATELLITES DANS L'IMAGE

Suite à l'étape de classification pixellaire (cf. chapitre 4), nous cherchons à déterminer le nombre de satellites LOS (resp. NLOS). Pour ce faire, nous projetons les satellites dans l'image et identifions les satellites se trouvant dans la région " ciel " (resp. dans la région " non ciel "). Nous détaillons ci-dessous la méthodologie utilisée pour placer les satellites dans l'image. Celle-ci se décompose en trois étapes : 1) l'acquisition des données ; 2) le calibrage et la mise en forme des données ; 3) le recalage des données.

- 1) Acquisition des données. Le dispositif d'acquisition est composé d'une caméra mono-CCD couleur et d'un GPS Septentrio (récepteur et antenne). La caméra, équipée d'un objectif fisheye, est fixée sur le toit d'un véhicule et orientée verticalement vers le ciel. La fréquence d'acquisition des images est de 27 images par seconde. Parallèlement à l'acquisition des images, les trames *NMEA* [ProFlex 2009], issues du GPS Septentrio, sont également enregistrées.
- 2) Calibrage et mise en forme des données images. Les images présentent un phénomène de distorsion. Cette aberration liée aux propriétés optiques des lentilles de l'objectif fisheye se caractérise par une déformation de l'image. Un calibrage, basé sur une mise en correspondance des points du monde réel avec les pixels de l'image est nécessaire pour corriger cette déformation. Il est en effet indispensable de positionner les satellites dans chaque image avec une précision de l'ordre du pixel. Deux pistes ont été suivies par DAVID NOWAK (stagiaire sous l'encadrement d'AMAURY FLANCQUART) au LEOST de l'IFSTTAR [Nowak 2005], pour atteindre nos objectifs. La première s'appuie sur une méthode mathématique intégrant le modèle géométrique d'une caméra et ses paramètres intrinsèques. Cette méthode, basée sur l'algorithme de Zhai implémenté dans OpenCV, a montré ses limites pour un objectif offrant un champ de vision large (c'est le cas d'une caméra avec un objectif fisheye). La seconde méthode est en partie expérimentale. Le dispositif, mis en place en laboratoire et basé sur un système de mires (cf. figure 5.4), a permis de répondre à nos attentes avec la précision escomptée.
- 3) Enfin, les trames *GPGGA* et *GPGSV* extraites des données *NMEA*, issues du GPS Septentrio, nous informent sur la position du véhicule toutes les secondes, ainsi que sur les positions des satellites reçus en termes d'azimut et d'élévation. Les

informations temporelles associées à chaque système d'acquisition nous permettent de recalibrer les données de localisation, les images, puis de placer correctement les satellites dans l'image.

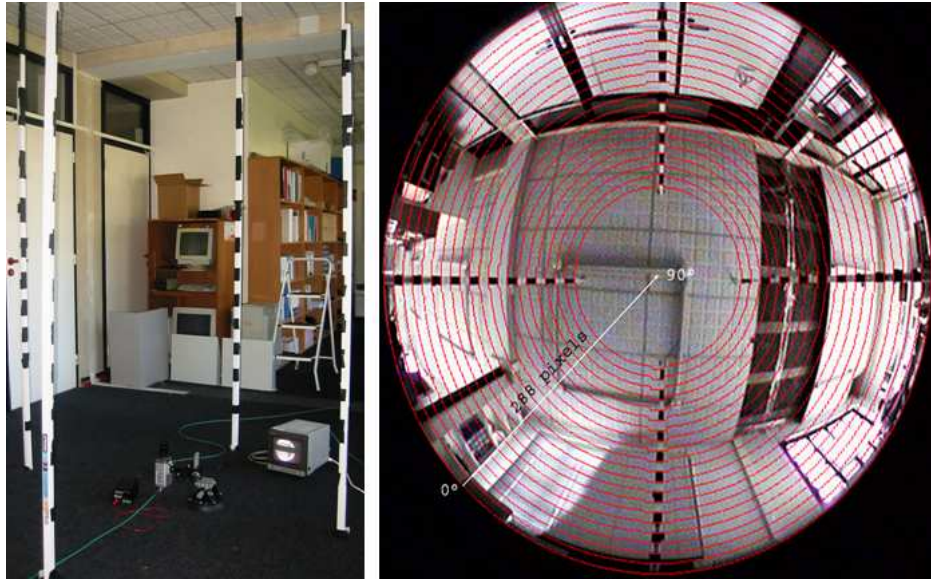


FIGURE 5.4 – Dispositif de calibrage [Nowak 2005].

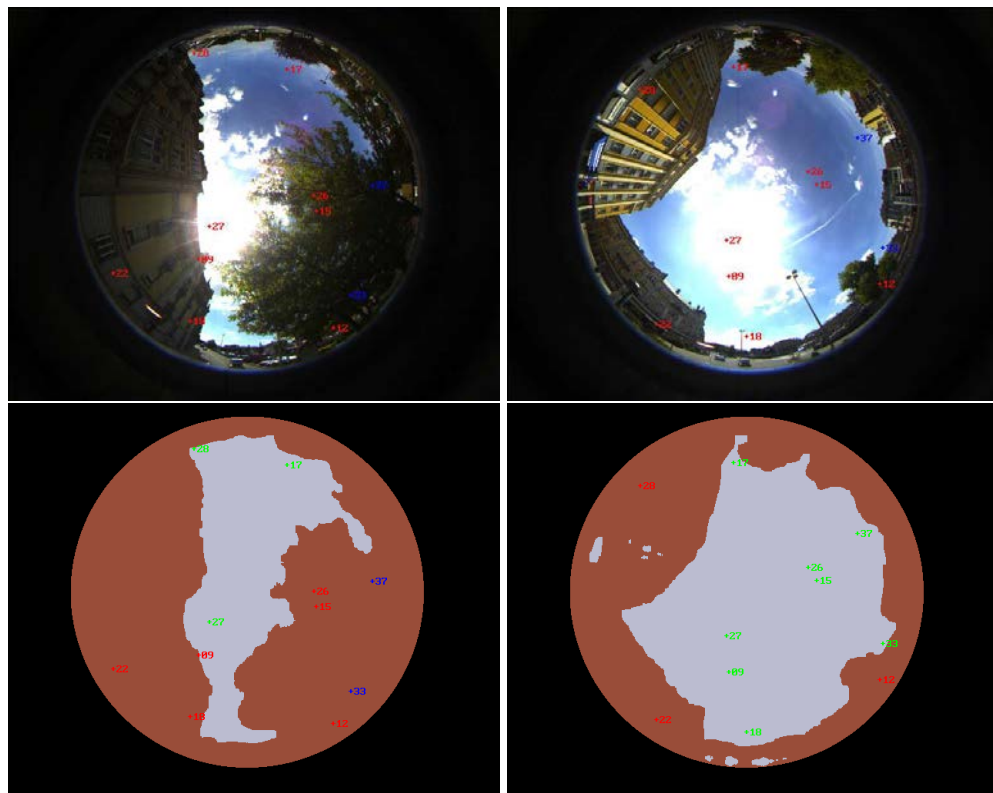


FIGURE 5.5 – Illustration de la détection du ciel avec repositionnement des satellites dans l'image (de haut en bas : images initiales, images classées par l'algorithme Fisher après filtrage médian).

La figure 5.5 illustre la projection des satellites sur deux images de la base CAPLOC. Les satellites situés dans une région "ciel" sont représentés par une couleur verte, tandis que les satellites situés dans une région "non ciel" sont représentés par une couleur rouge (pour la constellation GPS) ou bleue (pour la constellation EGNOS). Nous pouvons ainsi identifier directement, les satellites LOS dont les signaux suivent un trajet direct (en vert) et les NLOS dont les signaux sont bloqués ou réfléchis (en rouge ou bleu).

#### 5.4.2/ DÉTECTION DE LA LIGNE D'HORIZON

La ligne d'horizon est la frontière qui sépare les régions "ciel" et "non ciel". Tout ce qui constitue un obstacle à la réception des signaux est considéré comme "non ciel" (végétation, bâti, véhicules, etc.).

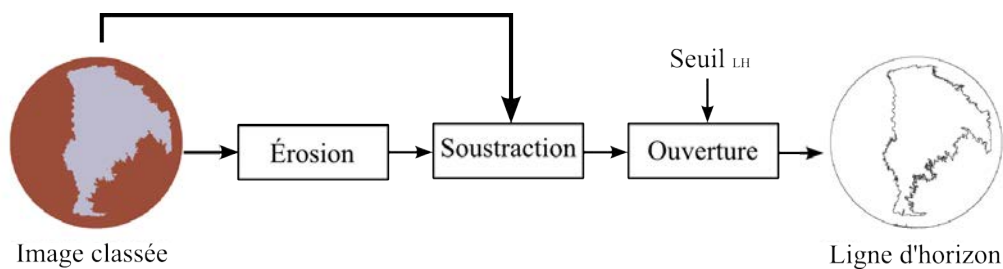


FIGURE 5.6 – Principe de la détection de la ligne d'horizon.

La détection de la ligne d'horizon, réalisée à partir de l'image classée, est principalement basée sur des opérateurs issus de la morphologie mathématique. La figure 5.6 décrit les étapes de cette opération. La première étape consiste à appliquer une érosion de l'image classée. Ensuite, une soustraction entre l'image classée et l'image érodée permet de trouver l'ensemble des pixels séparant les deux classes "ciel" et "non ciel". La dernière étape est une étape de post-traitement. Elle consiste à appliquer, sur le résultat de soustraction une ouverture morphologique, et à supprimer ainsi les ensembles de pixels isolés ayant un cardinal inférieur à un seuil  $Seuil_{LH}$ . Ce dernier ne doit pas être supérieur à la longueur de la ligne d'horizon. Le seuil a été choisi empiriquement, dans notre cas  $Seuil_{LH} = 500$  pixels. La figure 5.7 illustre, sur deux images de la base CAPLOC, la détection de la ligne d'horizon, affichée en couleur verte.

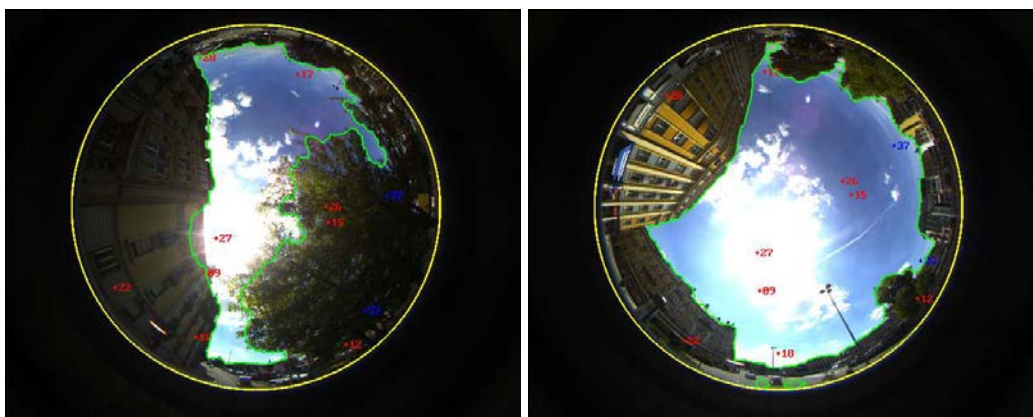


FIGURE 5.7 – Détection d'une ligne d'horizon.



## 5.5/ CALCUL D'UNE NOUVELLE POSITION

Dans cette section, nous détaillons les techniques que nous proposons pour estimer une position plus précise. Nous allons commencer par appliquer la stratégie d'exclusion des satellites NLOS. Ensuite, nous introduisons l'information DOP caractérisant la géométrie des satellites pour pouvoir choisir le moment approprié où l'application de l'exclusion des satellites NLOS améliore la précision de la position estimée.

### 5.5.1/ SYSTÈMES D'ACQUISITION

Nous rappelons le lecteur que dans le cadre du projet CAPLOC, le système d'acquisition des données est composé de différents capteurs vidéo et GNSS.

- Un récepteur GPS classique de marque uBlox, implanté dans un équipement Safedrive.
- Un récepteur GPS/EGNOS de marque Septentrio, plus haut en gamme.
- Un récepteur GNSS-RTK. Le GPS RTK est un système de positionnement précis. Nous l'utilisons comme source de positions de référence.
- Une caméra mono-CCD couleur équipée d'un objectif grand angle (fisheye).

Ce système complet a été installé sur un véhicule de laboratoire pour réaliser des campagnes d'acquisition des données en mobilité. La disposition de ces équipements est illustrée par la figure 4.2 du chapitre 4, décrivant les véhicules de l'UTBM/IRTES-SET et celui du laboratoire IFSTTAR-LEOST.

### 5.5.2/ DESCRIPTION DES DONNÉES EXPÉRIMENTALES

Une campagne d'acquisition réalisée en juin 2010 avec les véhicules en mobilité, a permis d'enregistrer un total de 1885 positions estimées, soit une durée de 31 minutes et 25 secondes, à raison d'une position par seconde. Nous nous sommes intéressés, dans cette expérimentation, à l'intervalle de temps [13 : 15 : 46; 13 : 16 : 56] soit 71 données GPS. L'enregistrement des images a été effectué à  $27Hz$  et les trames *NMEA* sont cadencées à  $1Hz$ . Plusieurs images correspondent donc à une donnée GPS. Afin de réduire le temps de traitement, un échantillonnage de la séquence d'images a été effectué afin de ne garder qu'une image sur 10 (la première de chaque dizaine). Au final, à chaque seconde, la base de données est constituée de 2 à 3 images. Cet échantillonnage est justifié par la faible variation de la répartition des satellites entre les images successives acquises pendant une seconde.

La figure 5.8 illustre la trajectoire étudiée. La trajectoire en rouge est formée par les positions de référence obtenues à l'aide du récepteur GPS-RTK. La fréquence d'acquisition de ce dernier est plus importante que celle des autres récepteurs. Nous avons exploité une portion de trajectoire composée de 171 points. La trajectoire en bleue est constituée des positions estimées à partir des signaux reçus depuis tous les satellites disponibles, en utilisant une technique basée sur le filtre de KALMAN étendu implémenté par VIANDIER [Viandier 2011]. Les pseudo-distances sont fournies par l'équipement Safedrive et utilisées dans le filtre. La totalité de la trajectoire est formée de deux portions quasiment perpendiculaires et reliées par un virage. La première ligne droite du trajet est longue d'environ 81 mètres et comprend 32 points de mesure. Parcourue avec une



vitesse moyenne inférieure à  $10\text{km}$  par heure, cette ligne droite commence par un temps d'arrêt identifiable sous la forme d'un amas de points, que nous pouvons retrouver sur les figures 5.9 et 5.14. Sur la deuxième ligne droite, on compte 166 mètres et 26 points, avec une vitesse moyenne de parcours de  $23\text{km}$  par heure.

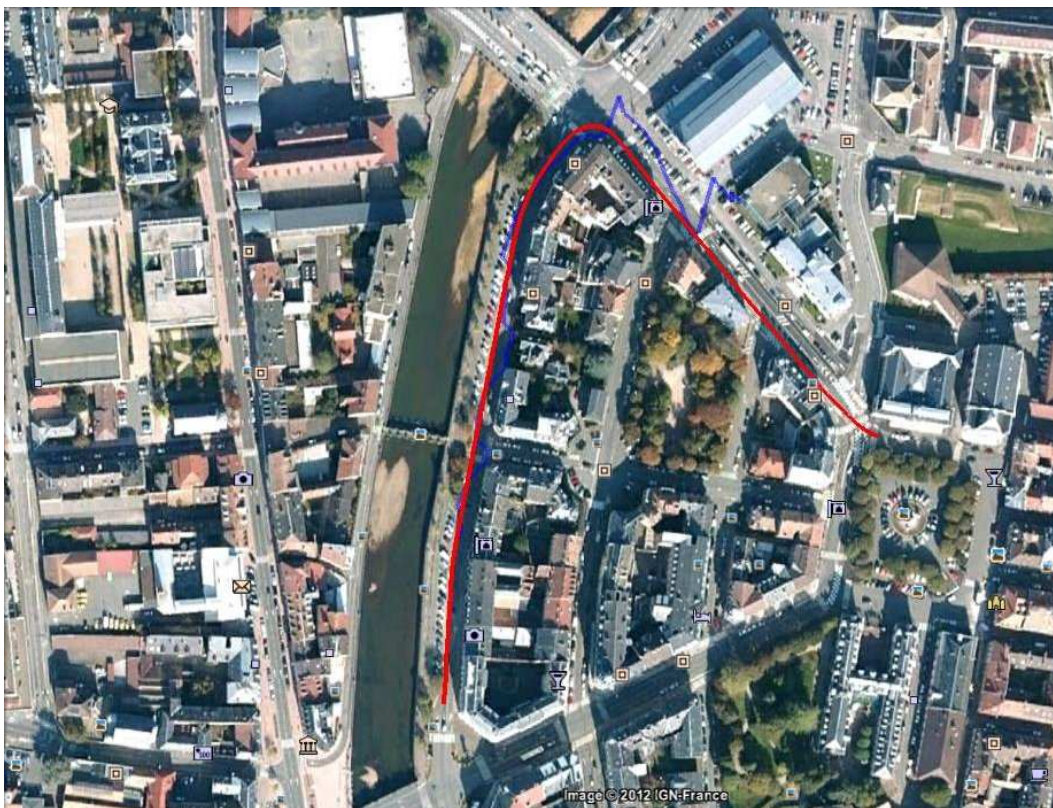
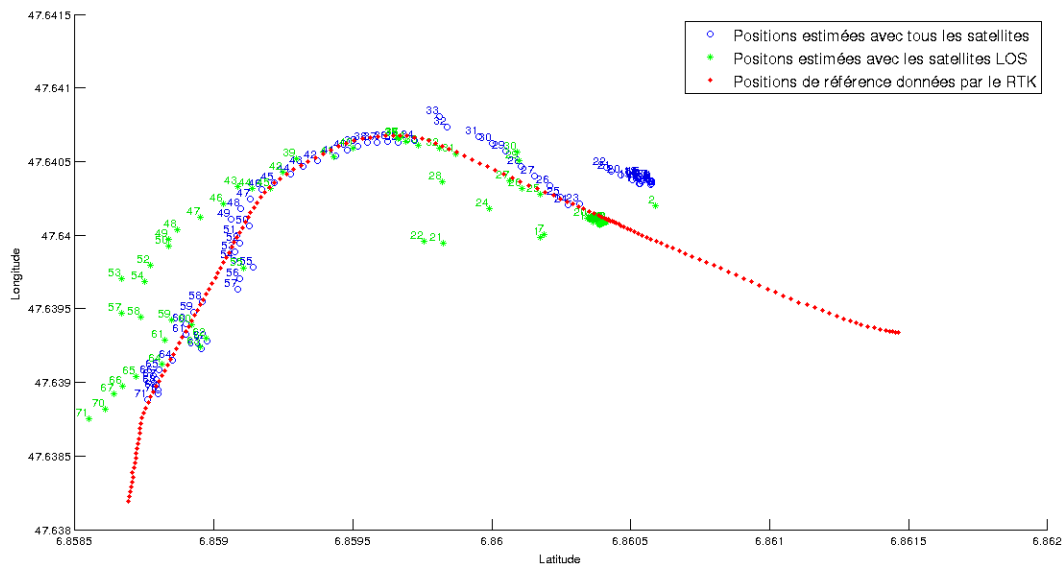


FIGURE 5.8 – Trajectoire étudiée : positions de référence (en rouge) et positions estimées (en bleue).

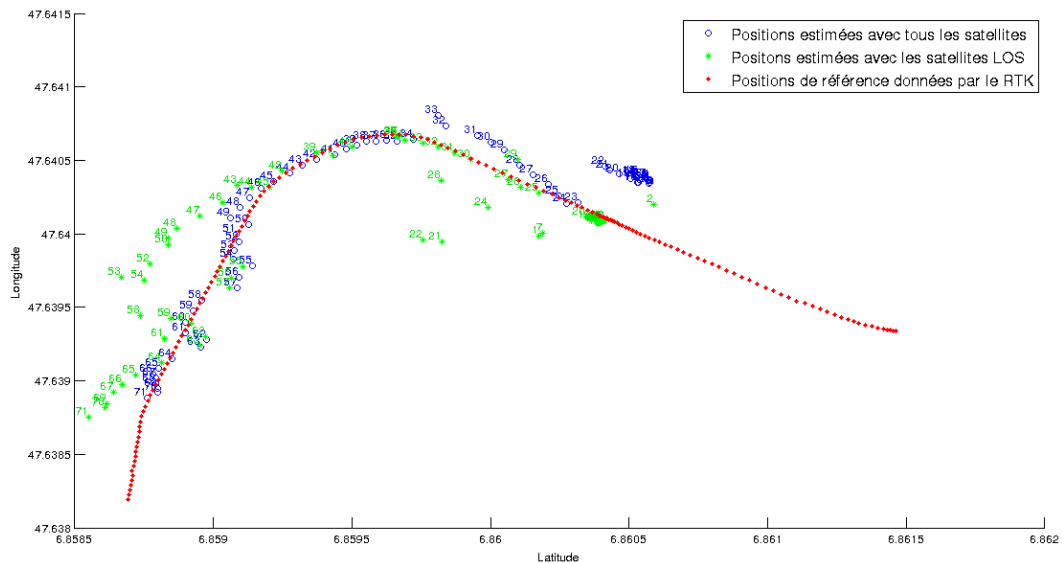
### 5.5.3/ CALCUL DE POSITION À PARTIR DES SATELLITES LOS UNIQUEMENT

Dans le but d'améliorer la position calculée à partir des signaux disponibles, nous appliquons la stratégie d'exclusion des satellites dont les mesures peuvent être considérées comme dégradées. Comme décrit précédemment, cette solution est basée sur l'analyse d'images acquises simultanément avec les données GPS (les images décrivant l'environnement de réception des signaux GNSS). Nous avons choisi d'exclure les satellites NLOS (se trouvant dans les régions " non ciel ") afin d'éliminer toute source mettant en jeu le phénomène de multi-trajet. À partir de ce critère de tri, nous avons recalculé la position en utilisant uniquement les signaux reçus par les satellites se trouvant dans la région " ciel ". Nous avons ensuite comparé les deux trajectoires obtenues en considérant tous les satellites (LOS et NLOS) d'une part et seulement les satellites LOS d'autre part.

Dans les figures 5.9 et 5.14, la trajectoire de référence, toujours tracée en rouge, permet d'évaluer le gain des différentes solutions testées. La courbe bleue représente un résultat de référence issu d'une solution classique (solution utilisant tous les satellites).



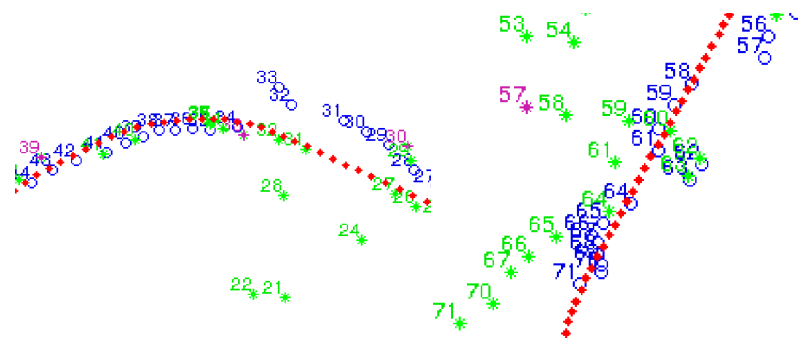
(a) Cas 1 : la première image est choisie à chaque instant.



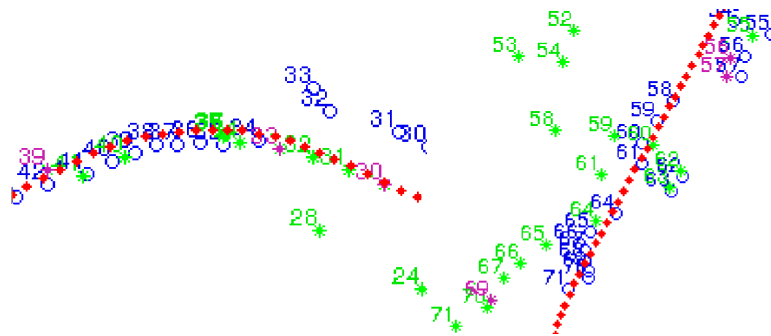
(b) Cas 2 : l'image présentant le nombre maximal de satellites LOS est choisie à chaque instant.

FIGURE 5.9 – Trajectoires obtenues en n'utilisant que les satellites LOS pour le calcul des positions.

Les figures 5.9(a) et 5.9(b) représentent les trajectoires calculées avec deux stratégies différentes pour le choix de l'image à associer à la mesure GPS. Comme décrit plus haut, plusieurs images peuvent correspondre à un même point GPS. Nous avons alors opté pour deux cas d'utilisation. Dans le premier cas (cf. figure 5.9(a)), le choix est arbitraire : la première image de chaque instant  $t$  est utilisée pour la classification des satellites. Dans le deuxième cas (cf. figure 5.9(b)), la stratégie est plus optimiste. Partant du principe que l'apparition d'un obstacle très furtif ne perturbera pas le signal, les 2 ou 3 images disponibles (après échantillonnage) à l'instant  $t$  sont traitées, et l'image offrant la meilleure disponibilité (ie. le plus grand nombre de satellites LOS) est conservée.



(a) Zoom sur les instants 30, 33, 39 à droite et sur l'instant 57 à gauche.  
Absence des instants 56 et 69 - Cas 1.



(b) Zoom sur les instants 30, 33, 39 à droite et sur l'instant 56 à gauche.  
Apparition des instants 57, 69 - Cas 2

FIGURE 5.10 – Illustration des différences entre les deux cas d'utilisation.

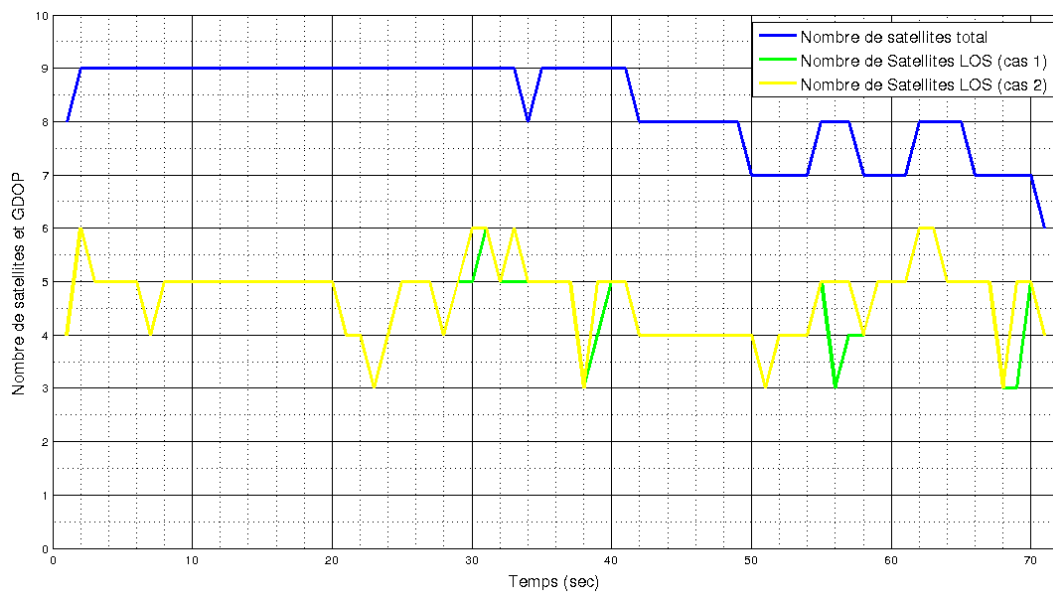


FIGURE 5.11 – L'évolution dans le temps du nombre total de satellites et du nombre de satellites LOS.

Les deux cas d'utilisation fournissent des résultats très proches. En effet, sur l'ensemble de la trajectoire, la précision obtenue avec les deux cas est la même (points identiques). Le deuxième cas permet cependant de réduire de quelques points l'indisponibilité de

service grâce au caractère optimiste de la stratégie de choix de l'image permettant la classification des satellites. En effet, un satellite pouvait ne pas être visible sur la première image mais l'être sur les autres images de la même seconde. L'évolution du nombre de satellites LOS dans le temps, représentée sur la figure 5.11 pour les deux cas, permet d'identifier clairement ces points (différence de disponibilité) aux instants 30, 33, 39, 56, 57 et 69. Nous pouvons noter aussi l'apparition de positions aux instants 56 et 69 grâce au passage de trois satellites LOS à 5 satellites LOS (cf. figure 5.10).

Nous observons alors qu'en éliminant les satellites NLOS, la précision est améliorée sur la première moitié du trajet ainsi que dans le virage où les positions estimées sont visiblement plus proches de la trajectoire de référence, fournies par les mesures GPS-RTK. L'apport de l'exclusion des satellites NLOS reste cependant limité sur la deuxième partie de la trajectoire (la fin de la courbe verte dans les figures 5.9(a) et 5.9(b)). Nous pouvons expliquer ceci par l'impact de la géométrie des satellites utilisés lors du calcul de la nouvelle position. Pour illustrer la relation entre la géométrie des satellites et la précision de la position estimée, nous utilisons le GDOP dont l'évolution est représentée sur la figure 5.12. Rappelons qu'une valeur de GDOP caractérisant une bonne géométrie est classiquement considérée inférieure à 3. À l'exception de quelques points de GDOP largement supérieur à 2, l'ensemble des points calculés avant  $t = 40s$  bénéficient d'une bonne géométrie, même lorsque les satellites NLOS sont exclus (points de couleur jaune et verte). La situation est clairement dégradée sur la période  $[40, 60]$  et l'impact sur la trajectoire est visible directement. Cette configuration montre les limites de notre stratégie : les satellites disponibles, même s'ils sont directs, ne sont pas disposés de façon optimale et la position est moins précise que lorsque tous les satellites sont utilisés.

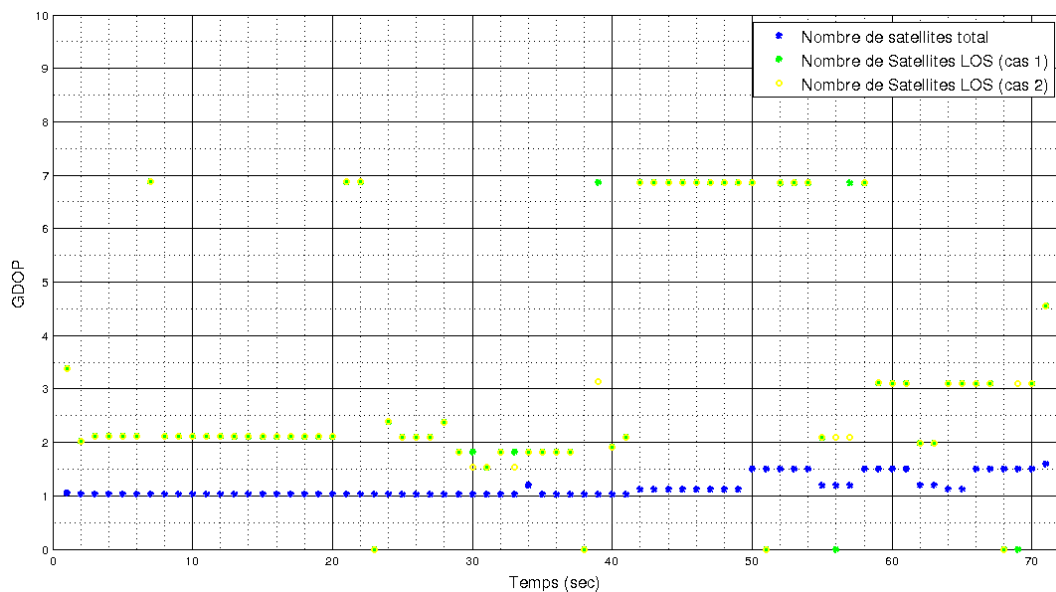


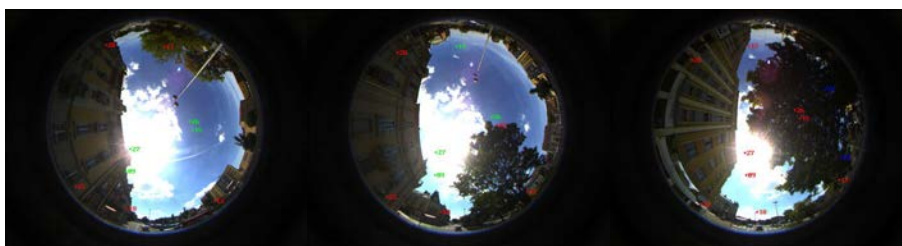
FIGURE 5.12 – L'évolution dans le temps du GDOP des satellites LOS (cas 1 et 2) et de tous les satellites.

Si l'on observe la configuration des environnements sur les deux moitiés de la trajectoire, on constate que la première moitié de la trajectoire est orientée vers le nord-ouest avec des obstacles symétriques des deux côtés (cf. figure 5.13(a)). Dans cette première partie,

le masque laisse des satellites visibles de part et d'autre de l'antenne de réception. La deuxième partie du trajet, orientée vers le sud-ouest, est bordée à gauche par des immeubles et à droite par un quai longeant la rivière (La Savoureuse). Cette dernière est parfois masquée par de la végétation (cf. figure 5.13(b)). La géométrie du masque est ici très déséquilibrée.



(a) Illustration d'images de la première partie du trajet.



(b) Illustration d'images de la deuxième partie du trajet.

FIGURE 5.13 – Illustration de la répartition géométrique des satellites sur les deux parties du trajet.

#### 5.5.4/ UTILISATION COMBINÉE DU DOP ET DE L'EXCLUSION DES SATELLITES NLOS

Au vu des résultats et des discussions notés plus haut, et pour optimiser notre démarche de calcul d'une nouvelle position, l'approche choisie consiste à combiner l'information DOP et la stratégie d'exclusion des satellites NLOS. Cette combinaison consiste à appliquer la stratégie d'exclusion des satellites NLOS uniquement lorsque le DOP est suffisamment bon (valeur inférieure à un seuil choisi empiriquement - ici 3) et à garder tous les satellites dans le cas contraire.

La figure 5.14 montre l'apport de la combinaison du DOP et de la stratégie d'exclusion des satellites NLOS. La figure 5.15 illustre les résultats de la figure 5.14 à l'aide de l'outil Google Earth. L'approche par combinaison (exclusion des satellites NLOS conditionnée par la valeur du GDOP) montre bien l'exploitation de l'information la plus avantageuse dans chaque situation.



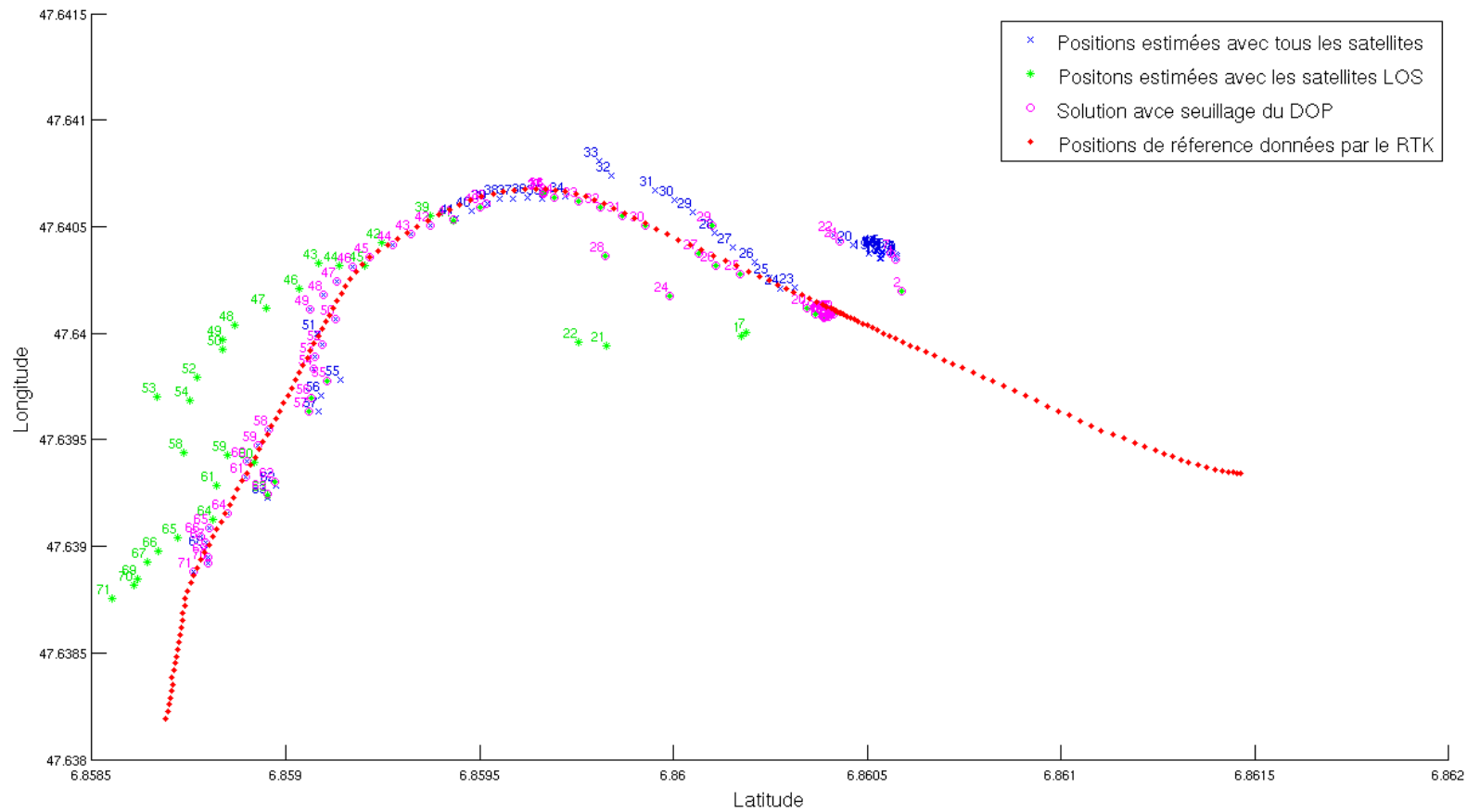


FIGURE 5.14 – Comparaison entre les trois solutions de calcul de position : la trajectoire rouge est la référence (GPS-RTK), la trajectoire bleue issue de la solution utilise tous les satellites, la trajectoire verte est issue de la solution qui n'utilise que les satellites LOS, la trajectoire violette est issue de la solution par combinaison (exclusion des satellites NLOS si la valeur du GDOP est inférieur au seuil choisi ( $Seuil_{GDOP} = 3$ )).



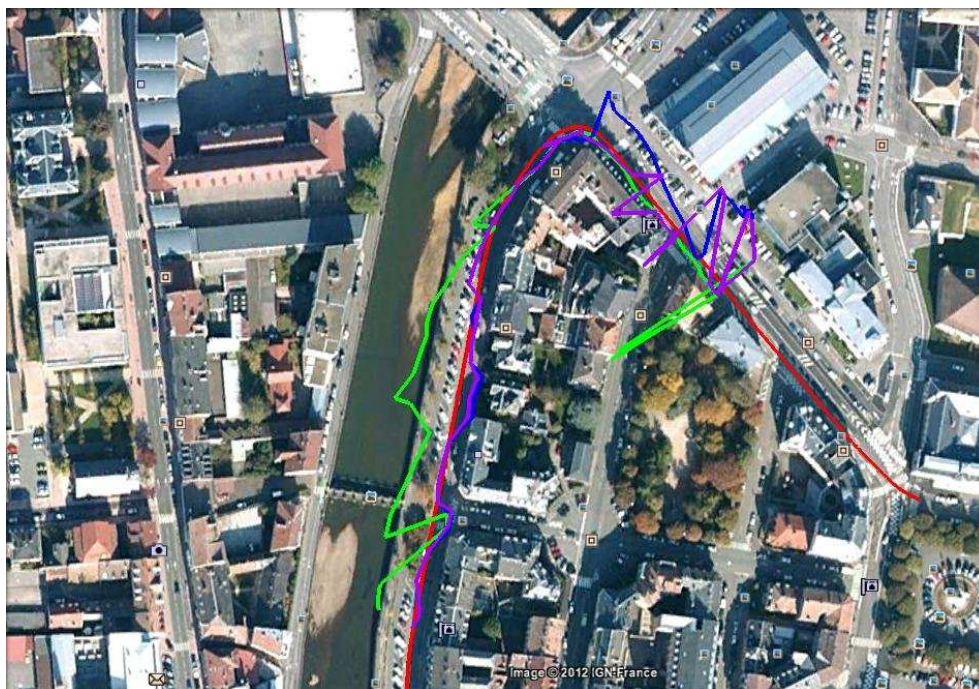


FIGURE 5.15 – Vue aérienne (Google Earth) des trajectoires obtenues avec les trois solutions (en rouge : la référence (GPS-RTK), en bleu : tous les satellites sont utilisés, en vert : les satellites LOS sont uniquement utilisés, en violet : la solution par combinaison).

#### 5.5.5/ DISCUSSIONS SUR L'IMPACT DE LA VÉGÉTATION SUR LA VISIBILITÉ SATELLITAIRE

La littérature recense quelques travaux traitant l'impact de la forêt sur la précision des récepteurs GPS [Bettinger 2012]. Ces travaux s'adressent surtout à la communauté des utilisateurs forestiers et ne traitent pas directement des retards induits sur les pseudo-distances. Les articles proposant des modèles de canal de propagation GNSS s'intéressent à la modélisation de la propagation à travers la végétation : arbres et allées forestières. Les modèles développés et les conclusions qui y sont attachées sont difficiles à généraliser en raison de la diversité des espèces et saisons, en raison de la variation des effets d'une situation à une autre. Les effets de propagation dans les arbres sont : le masquage et l'atténuation (par les feuilles), la diffusion (par les branches), la réflexion et la diffraction (par le tronc) [Schubert 2010]. L'impact est important lorsque plusieurs arbres sont regroupés, et moins visible lorsque le masquage est le fait d'un seul arbre isolé. Selon SCHUBERT, les arbres situés le long de la route atténuent périodiquement le signal de  $-10$  à  $-30$  dB quand le signal LOS est masqué. L'influence de la diffusion peut être importante, mais intervient lorsque le satellite est LOS. Ce cas n'est pas traité dans cette thèse. Selon SCHUBERT [Schubert 2012], l'influence du tronc est négligeable.

Rappelons que pour nos premiers traitements appliqués sur l'image, seulement deux classes ont été retenues ("ciel" et "non ciel"). La végétation a été considérée comme un masque aussi bloquant pour les signaux que les bâtiments ou autres obstacles. Dans cette section, cette hypothèse est discutée.

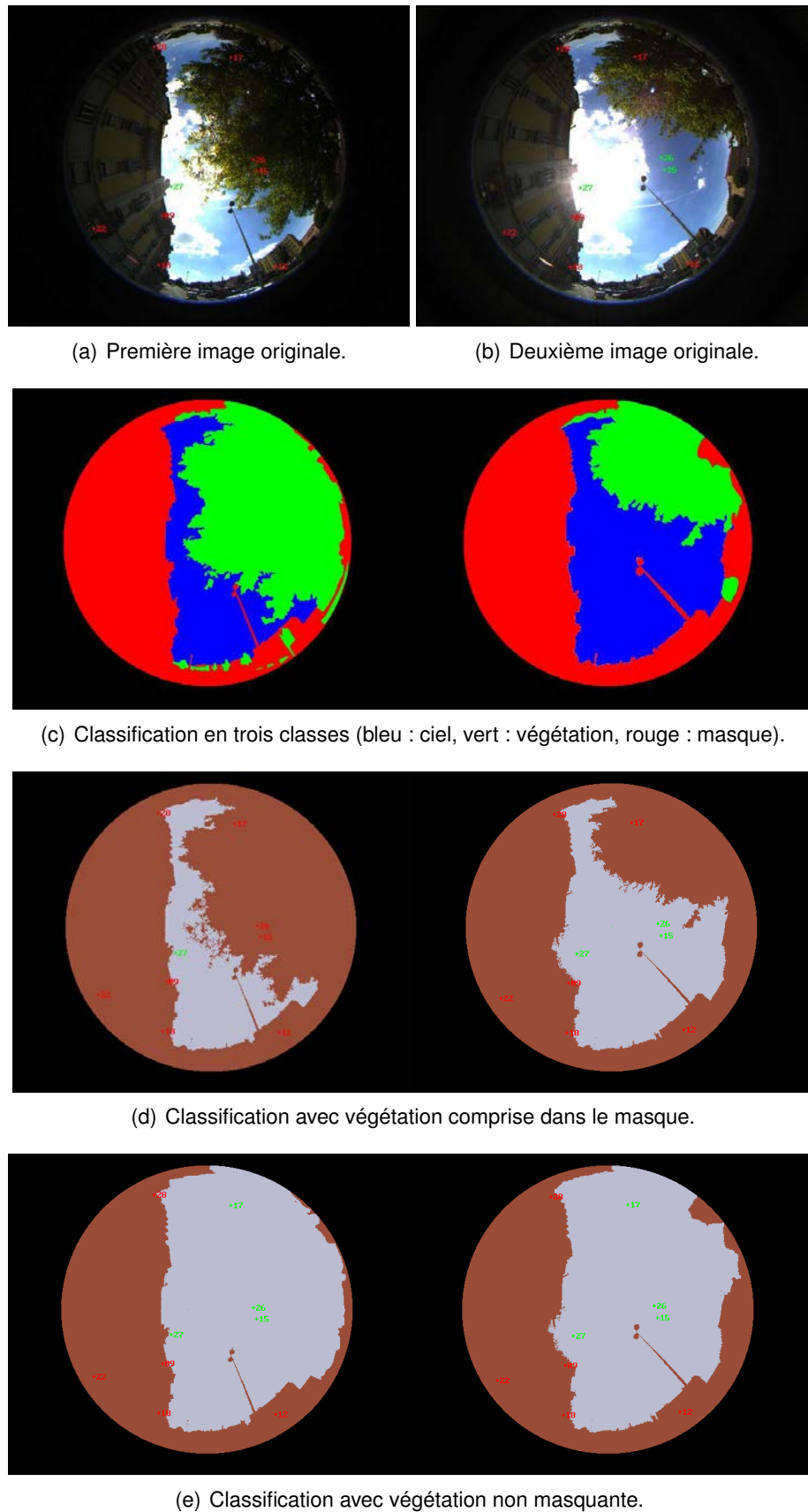


FIGURE 5.16 – Exemple, sur deux images issues de la base du projet CAPLOC, de changement de l'état de satellites de NLOS à LOS, en fonction de la classification de la végétation en masque ou non.

Dans ce qui suit, la végétation est considérée à part. Si l'on reprend une partie des conclusions de SCHUBERT [Schubert 2009], nous pouvons considérer, d'un point de vue propagation, que l'amplitude des signaux issus des satellites placés dans la végétation est assez faible pour qu'une partie des signaux puissent être exploitée par le récepteur. Dans notre classification des états de réception, un satellite placé dans la végétation sera classé comme LOS. Ici, la procédure est effectuée manuellement (le nombre de satellites placés dans la végétation étant faible) et vise à déterminer si une classification en trois classes (ciel, masque et végétation) pourrait être une perspective de nos travaux.

La figure 5.16 illustre, sur deux images, un exemple de changement de l'état des satellites NLOS en LOS, s'ils se trouvent dans la végétation. Dans l'image 5.16(a), les satellites NLOS numéros 15, 17 et 26 sont reclassés en LOS avec la nouvelle hypothèse (cf. image 5.16(c)). En ce point précis, cette modification permet de ramener une disponibilité de service en faisant augmenter le nombre de satellites utilisés de 1 à 4. Sur la trajectoire de test, cela correspond aux deux instants ( $t = 38s$  et  $t = 68s$ ) (cf. figure 5.17).

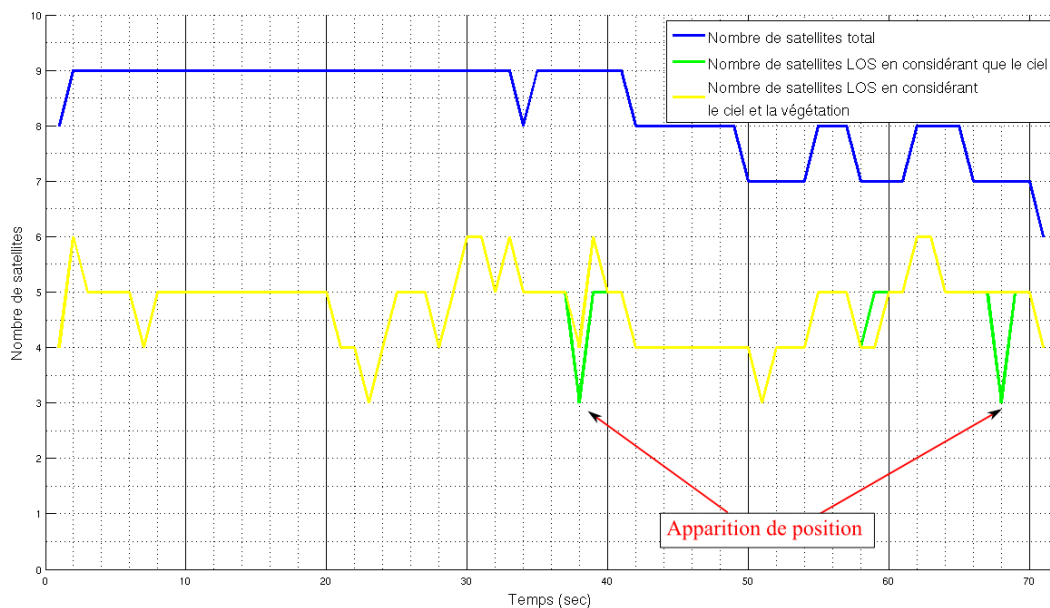


FIGURE 5.17 – Comparaison entre les deux hypothèses (LOS  $\equiv$  Ciel et LOS  $\equiv$  Ciel-Végétation).

## 5.5.6/ ANALYSE QUANTITATIVE

Afin d'évaluer quantitativement les performances des solutions de calcul de position proposées, il convient de calculer l'erreur de mesure en chaque point.

### 5.5.6.1/ ESTIMATION DE LA PRÉCISION D'UNE POSITION CALCULÉE

Plusieurs critères pour estimer la précision d'une position calculée peuvent être utilisés. VIANDIER mentionne, dans ses travaux [Viandier 2011], trois critères :

- 1) Le premier critère (exprimé en mètres) est la précision qui consiste à calculer la différence entre la position estimée et la vraie position.
- 2) Le deuxième critère (exprimé en pourcentage) est la continuité de service qui consiste à analyser la durée pendant laquelle le service est utilisable avec une qualité suffisante. La qualité peut être une précision minimale ou un niveau d'intégrité.
- 3) Le troisième critère relève des critères de "sûreté de fonctionnement", parmi lesquels trois mesures sont exprimées en secondes : le MUT (Mean UP Time - temps moyen de fonctionnement), le MDT (Mean Down Time - temps moyen de dysfonctionnement) et MTBF (Mean Time Between Failure - temps moyen entre deux défaillances) [Beugin 2009].

S'agissant de mesures statistiques, ces deux derniers critères sont judicieux lorsque la quantité de données est suffisamment représentative. Dans notre application, et en raison de la courte longueur de la trajectoire, seul le critère de précision a été retenu.

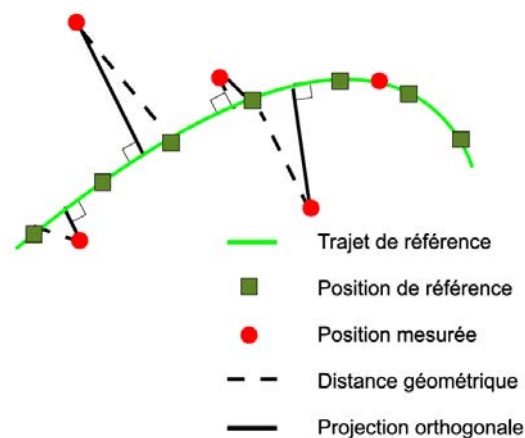


FIGURE 5.18 – Principe de la projection orthogonale pour le calcul de l'erreur de position [Viandier 2011].

La précision est calculée par rapport à la trajectoire de référence obtenue avec le récepteur GPS-RTK. Dans notre première campagne d'acquisition, les données acquises (images et GNSS) n'ont pas été synchrones. Comme la fréquence des points fournis par le GPS-RTK est importante, une projection orthogonale de chaque point sur la trajectoire a été appliquée pour estimer cette précision. Cette estimation n'est effectuée que lorsque la position peut être calculée (c'est-à-dire s'il y a suffisamment de satellites LOS). La projection orthogonale consiste à calculer, comme représenté sur la figure 5.18, la distance minimale entre la position mesurée et le trajet de référence.

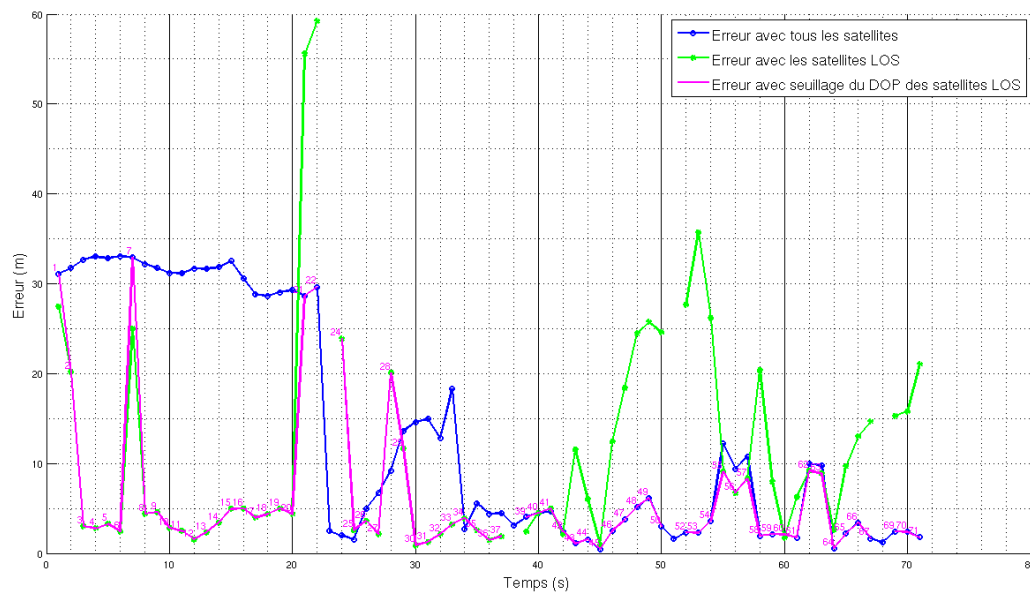


FIGURE 5.19 – Courbes d'erreur en fonction de la position estimée avec chacune des trois solutions d'estimation de position.

La figure 5.19 représente, en fonction du temps, les courbes d'erreur calculées pour les trois solutions proposées : avec tous les satellites, avec les satellites LOS et avec la solution par combinaison (exclusion des satellites NLOS conditionnée par le seuillage du DOP). La figure 5.19 montre que la solution basée sur le seuillage du DOP (approche par combinaison) est à chaque instant plus précise que les deux autres solutions (avec tous les satellites et avec le tri de satellites LOS) sauf en 3 points  $t = 1s$ ,  $7s$  et  $39s$ .

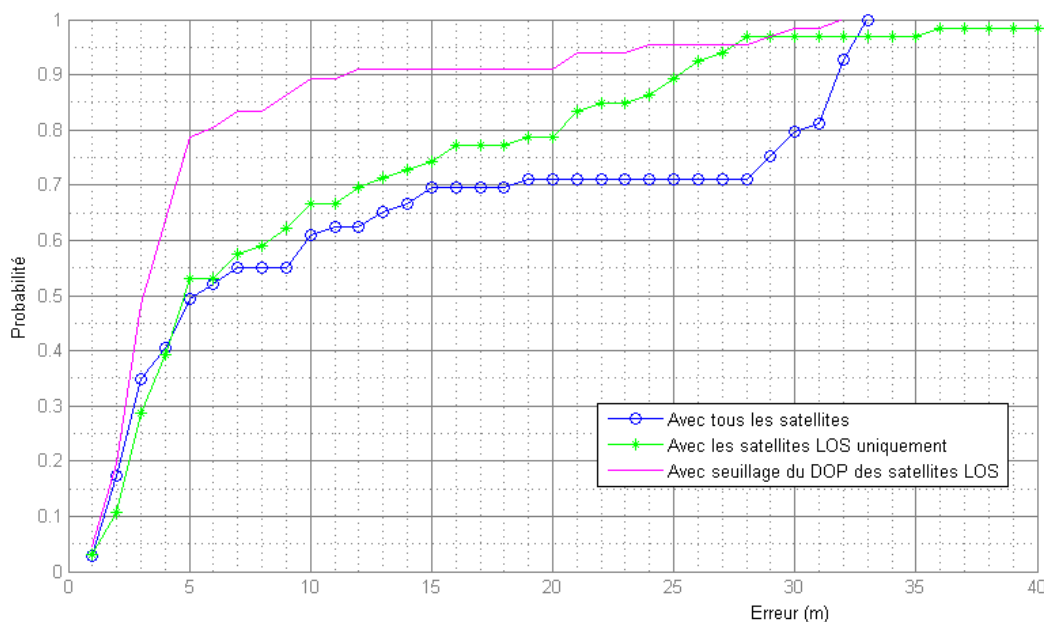


FIGURE 5.20 – Fonction de répartition de l'erreur de positionnement selon la solution utilisée.



Pour une meilleure visualisation, nous présentons sur la figure 5.20, les résultats de la figure 5.19, mais cette fois, sous la forme d'une fonction de répartition. Cette représentation met en évidence l'amélioration apportée par les différentes solutions sur la précision d'origine. Nous pouvons constater que la médiane de l'erreur passe de 5m avec la prise en compte de tous les satellites à 3m avec la solution par combinaison (choix final), soit une amélioration de 40%. Nous constatons aussi que 95% des positions sont inférieures à 24m avec l'approche d'exclusion des satellites NLOS conditionnée par le seuillage du DOP, contre seulement 70% des mesures avec la solution utilisant tous les satellites disponibles.

Stratégies		Mesures d'erreur (m)		
		Moyenne	Maximale	écart type
Tous les satellites		13.13	33.07	12.75
Satellites LOS	Ciel et 1 <sup>ère</sup> image	10.5	59.23	12.14
	Ciel et Max Sat	10.25		11.92
	Ciel-Végétation et 1 <sup>ère</sup> image	10.69		11.79
	Ciel-Végétation et Max Sat	10.7		11.98
Seuillage du DOP des satellites LOS	Ciel et 1 <sup>ère</sup> image	5.73	32.92	7.66
	Ciel et Max Sat	5.59		7.52
	Ciel-Végétation et 1 <sup>ère</sup> image	5.85		7.57
	Ciel-Végétation et Max Sat	5.6		7.5

TABLE 5.1 – Performances des différentes approches de calcul de position proposées.

Le tableau 5.1 reprend trois mesures calculées pour chacune des stratégies de calcul de position testées : l'erreur moyenne, l'erreur maximale et l'écart type. Les stratégies testées sont : l'utilisation de tous les satellites disponibles de la constellation, la sélection des satellites LOS uniquement et la solution par combinaison (seuillage du DOP des satellites LOS). Pour les deux dernières stratégies, on considère deux cas comme précédemment : utilisation de la première image à chaque instant après échantillonnage (avec ou sans la végétation masquante), et utilisation de l'image la plus optimiste en terme de disponibilité (avec ou sans la végétation masquante). Ces résultats montrent que la stratégie basée sur le seuillage du DOP est toujours meilleure, comparée aux deux autres stratégies (utilisation de tous les satellites et exclusion des satellites NLOS). La précision des positions estimées est améliorée d'un facteur 2 en utilisant la solution du seuillage du DOP. D'autre part, dans le cas de l'utilisation des satellites LOS, l'intégration des signaux reçus à travers la végétation conduit à une augmentation de l'erreur moyenne (cf. tableau 5.1). Dans le cas du seuillage du DOP, cela est beaucoup moins visible. Ceci nous permet de conclure qu'une classification automatique en deux classes ("ciel" et "non ciel") sans avoir à isoler/distinguer la végétation est suffisante à ce stade. Il conviendra, néanmoins, de vérifier cette conclusion sur une base de données plus importante.



## 5.6/ CONCLUSION

Dans ce chapitre, nous avons présenté comment sont exploités des résultats du chapitre précédent dans l'amélioration, en termes de précision, du calcul de la position d'un mobile dans un environnement urbain. Cette application, qui s'inscrit dans le cadre du projet CAPLOC, consiste à déterminer l'état de réception de chaque satellite, puis utiliser cette information dans le processus de calcul de la position. L'état de réception d'un satellite peut prendre deux valeurs. La première, LOS, correspond à un satellite dont le signal est reçu via un trajet direct. La deuxième valeur, NLOS, correspond à un satellite dont le signal est bloqué par les obstacles environnants l'antenne de réception GNSS ou reçu via un multi-trajet après réflexion sur un ou plusieurs obstacles de l'environnement. Pour déterminer l'état de réception des satellites, leur position est projetée sur l'image fisheye traitée par les algorithmes du chapitre 4. Ces derniers permettent de classer les pixels de l'image en deux classes : " ciel " et " non ciel ". Ainsi, les satellites LOS (respectivement NLOS) sont les satellites situés dans une " ciel " (respectivement " non ciel ").

Une fois l'état de réception des satellites connu, nous avons proposé deux stratégies de calcul de la position, puis nous les avons comparées à une solution classique [Viandier 2011] utilisant tous les satellites disponibles dans un processus de calcul de position basé sur un filtrage de Kalman étendu (le même processus est utilisé dans les stratégies proposées). La première stratégie proposée est basée sur le principe d'exclusion de tous les satellites NLOS dans le processus de calcul de la position. L'évaluation de cette stratégie a montré l'intérêt du principe d'exclusion des satellites NLOS, notamment lorsque la géométrie des satellites LOS est favorable. Dans le cas contraire, le principe d'exclusion peut fournir une position moins précise. Ce principe peut même causer l'impossibilité de fournir une position si le nombre de satellites LOS est insuffisant pour procéder à la trilatération (il faut au moins 4 satellites). Face à cette limitation, nous avons proposé une deuxième stratégie basée sur la combinaison sur le principe d'exclusion des satellites NLOS et l'information DOP qui caractérise la répartition des satellites dans l'espace. Cette stratégie consiste à conditionner l'application du principe d'exclusion des satellites NLOS par la valeur du DOP. Plus précisément, il s'agit d'appliquer le principe d'exclusion des satellites NLOS dans le processus de calcul de position uniquement lorsque le DOP des satellites LOS est inférieur à un seuil donné, correspondant à une répartition géométrique favorable. Les tests effectués ont montré que la deuxième stratégie est plus adaptée.

Dans cette thèse, la problématique de notre application a été abordée avec le principe d'exclusion des satellites NLOS du processus de calcul de position. D'autres pistes [Peyret 2011] et [Bourdeau 2012] visent à utiliser tous les signaux des satellites, en estimant les biais induits par les satellites NLOS à partir de modèles 3D de l'environnement de réception. Une pondération de la contribution des signaux des satellites dans le processus de calcul de position peut être une piste de travail intéressante, avec la mise en œuvre de techniques de résolution basées sur les moindres carrés pondérés, au lieu du filtre de Kalman.

Nous avons montré que la vision par ordinateur, notamment le traitement d'images,

contribue de manière pertinente dans la résolution de la problématique posée. La détection automatique des textures, notamment pour la détection de la végétation devrait permettre d'approfondir, en s'appuyant sur des bases d'images plus importantes en taille et en complexité, l'hypothèse «la réception d'un signal GNSS à travers la végétation dégrade légèrement la mesure de la position», étudiée de manière préliminaire dans ce chapitre. Par ailleurs, la détection et la caractérisation de régions dans les images intéressent de nombreux industriels des télécoms et/ou de localisation. En effet, dans le cadre du projet SATLOC par exemple, les chercheurs font appel à des techniques de traitement et analyse d'images pour intégrer, dans des simulateurs de signaux GNSS (SPIRENT l'un des fournisseurs de ces simulateurs), des masques réalistes afin d'évaluer les performances de récepteurs GNSS destinés aux applications ferroviaires.



# CONCLUSION GÉNÉRALE

Dans ce manuscrit, nous avons présenté notre contribution dans le domaine de la segmentation d'images par combinaison des informations de couleur et de texture. Notre principal objectif a consisté à développer une méthode de combinaison adaptative couleur/texture et non paramétrique. Cette méthode permet de combiner un gradient couleur (ou plus) et un gradient texture (ou plus) pour ensuite générer un gradient structurel, utilisé comme image de potentiel dans l'algorithme de segmentation par Ligne de Partage des Eaux (LPE). L'approche proposée a été appliquée sur différentes bases de données différentes : une base générique, Berkeley Segmentation Dataset and Benchmark, qui concerne la segmentation d'images couleur, une base d'images de texture connue sous le nom VISTEX, une base applicative composée d'images réelles acquises en mobilité dans le cadre de travaux sur l'amélioration de la localisation de véhicules en milieu urbain. La dernière base est liée au volet applicatif de cette thèse. Il porte sur la caractérisation, en temps réel, de l'environnement de réception des signaux satellitaires dans le but d'améliorer la localisation d'un mobile en milieu urbain dense. Le calcul de la nouvelle position passe par l'exclusion des satellites détectés (par traitement d'images) dans une zone de " non ciel " (les signaux émis par ces satellites sont réfléchis par des obstacles environnants ou totalement bloqués). Les données d'expérimentation sont des données réelles acquises en mobilité à l'aide de deux véhicules de laboratoire instrumentés de récepteurs GNSS et d'une caméra couleur munie d'un objectif fisheye et orientée vers le ciel.

Nous avons tout d'abord effectué un large état de l'art des différentes méthodes de segmentation d'images par utilisation conjointe de la couleur et de la texture. Cette étape nous a permis de relever trois principales catégories de méthodes à savoir l'intégration implicite (où l'information texture et l'information couleur sont mutuellement dépendantes), l'intégration successive (où l'extraction de la couleur et de la texture est assurée via un processus en série) et l'intégration que nous définissons d'hybride, qui consiste à combiner l'information couleur et l'information texture lors du processus de segmentation. En s'inscrivant dans le cadre de la dernière catégorie de méthodes, nous avons développé une nouvelle approche hybride, adaptative et non paramétrique permettant la combinaison des informations de couleur et de texture. La méthode proposée est basée sur l'étude des axes principaux d'un ensemble de points 3D formés par des informations de couleur et de texture (extraites des principaux attributs couleur et texture de la littérature). La nouveauté de cette technique réside dans la définition d'un gradient structurel à l'aide de la combinaison hybride et non paramétrique de gradients couleur et texture (définis comme étant des descripteurs couleur et texture). Contrairement à d'autres méthodes de combinaison issues de la littérature, la combinaison couleur/texture proposée s'adapte localement au contenu de l'image et ne requiert aucun paramètre de pondération pour effectuer la combinaison. Afin de comparer la performance de notre méthode de combinaison (utilisée dans l'algorithme de LPE), face à la diversité des méthodes de segmentation de la littérature, il était

important de disposer d'une méthode d'évaluation adaptée. Ainsi, l'état de l'art a été étendu aux méthodes d'évaluation utilisées en segmentation d'images. Nous avons retenu un ensemble de métriques d'évaluation supervisées et non supervisées. Une étude comparative détaillée entre les résultats de segmentation basée sur la combinaison couleur/texture proposée et les résultats des principales méthodes de segmentation citées dans la littérature a été alors réalisée. Les tests ont été effectués sur deux bases d'images communément utilisées par la communauté de chercheurs en vision par ordinateur, à savoir la base d'images couleur générique de BERKELEY [Arbelaez 2011] (où chacune des images possède au moins cinq segmentations de référence) et la base de texture VISTEX [VisTex 1995] (dédiée aux images texturées et ne proposant pas d'images de référence). L'algorithme de segmentation LPE, utilisant le gradient structurel obtenu à l'aide de la combinaison couleur/texture proposée et une image de germes obtenue à l'aide d'une réduction adéquate de germes, a permis d'obtenir, dans l'espace  $(R, G, B)$ , des résultats de segmentation très satisfaisants sur la base de BERKELEY et acceptables sur la base de texture VISTEX. En effet, sur la base VISTEX, la position de la méthode de segmentation basée sur la combinaison proposée oscille entre la 4<sup>ème</sup>/11 et 9<sup>ème</sup>/11 selon la métrique utilisée. Sur la base d'images couleur de BERKELEY, la méthode de segmentation basée sur la combinaison proposée se situe en moyenne en 6<sup>ème</sup> position sur 12 méthodes testées. De plus, lorsque la méthode proposée n'occupe pas la première ou la deuxième position, l'écart numérique par rapport aux méthodes de segmentation qui la précèdent reste relativement faible. Nous pouvons dire que les meilleurs résultats de segmentation, utilisant la combinaison proposée, sont obtenus avec le gradient structurel résultant de la combinaison d'un gradient morphologique couleur et d'un gradient de texture basé sur le deuxième paramètre d'HARALICK  $H_2$  (le contraste), avec une réduction de germes de 2%.

Le travail réalisé dans le volet applicatif de cette thèse porte sur le développement d'un système automatique et temps réel de traitement d'images permettant de caractériser l'environnement de réception des signaux GNSS. À partir de données images acquises depuis une caméra installée sur le toit du véhicule en mobilité et orientée vers le ciel, nous avons cherché à détecter la ligne d'horizon séparant la zone masquante " non ciel " de la zone non masquante " ciel " de l'image. Pour ce faire, nous avons proposé deux approches. La première consiste à segmenter l'image par l'algorithme de segmentation LPE, en utilisant notre nouvelle combinaison couleur/texture basée sur l'analyse en composantes principales, suivie d'une classification adaptée permettant de classer les régions obtenues en deux classes d'intérêt (" ciel " et " non ciel "). Bien que cette démarche offre les résultats souhaités en terme de qualité de classification (avec un taux de bonne classification moyen de 95.7% sur la base d'images du projet LOCOPROL et de 97.9% sur la base d'images du projet CAPLOC), le temps de traitement global de l'approche est très important pour répondre à la contrainte temps réel imposée par l'application. En effet, 3 minutes et 40 secondes sont nécessaires pour traiter une image de taille  $288 \times 384$ , issue de la base d'images LOCOPROL, avec un PC Core (TM2) cadencé à 2.80GHz. La deuxième approche, orientée temps réel, est basée sur une simplification de l'image couplée à une classification pixellaire adaptée. Différentes méthodes de simplification issues de la littérature (filtre médian, filtre gaussien, reconstruction géodésique par dilatation, etc) ont été testées. De la même manière que pour la première approche, un certain nombre d'algorithmes de classification non-supervisés et supervisés (KMlocal, FuzzyCmeans, Fisher, SRM, Bayes, Kppv et SVM), reconnus pour leur performance, ont également été considérés.

Une première série de tests, effectuée sur les images de la base du projet LOCOPROL, a permis de déterminer les deux meilleures méthodes de simplification. Le filtre médian et la reconstruction géodésique par dilatation (RGD) offrent les meilleurs résultats, avec un taux de bonne classification de 95.7% obtenu avec un filtrage médian et une classification par Fisher sur la composante couleur bleue et un taux de bonne classification moyen de 96% obtenu avec une simplification par RGD et une classification par SVM. Une deuxième série de tests, effectuée sur des images (très hétérogènes et de complexité variée) de la base du projet CAPLOC, a permis de définir le meilleur couple " méthode de simplification/classifieur ". Nous avons ensuite constaté que les meilleurs taux de bonne classification sont obtenus en utilisant un filtre médian ou une simplification par RGD, couplé à une classification par l'algorithme de Fisher (sur la composante bleue) ou par l'algorithme des Kppv (avec un nombre de  $K_{ppv} = 6$ ). Une étape de post-traitement a été ajoutée au processus afin de réattribuer les pixels mal classés. Cette dernière étape nous a permis d'obtenir un taux de bonne classification moyen de 98.1% pour le couple " Médian/Fisher " (VS 97.5% avant post-traitement) et un taux de bonne classification moyen de 98.1% pour le couple " RGD/Kppv " (98% avant post-traitement). Enfin, la prise en compte de la contrainte temps réel imposée par l'application, nous a poussé à retenir le couple " Médian/Fisher " (0.035s VS 15.4s soit un rapport de 1/400). Ces conclusions seront validées sur un nombre important d'images lors de la prochaine campagne d'acquisition, prévue en fin d'année 2013, dans le cadre du projet CAPLOC.

La caractérisation de l'environnement de réception des signaux GNSS par l'analyse d'images est destinée à l'amélioration de la précision la position estimée d'un véhicule en mouvement dans un milieu urbain dense. La phase d'analyse d'images proposée nous a permis de distinguer deux classes de satellites : les satellites LOS - Line-Of-Sight, détectés dans la zone " ciel " (les signaux sont reçus via un trajet direct) et les satellites NLOS - Non-line-Of-Sight, détectés dans la zone " non ciel ", (signaux bloqués ou subissant des atténuations trop fortes pour être exploités à cause du phénomène de multi-trajet). Dans le cadre de cette thèse, l'estimation de la position du récepteur a été effectuée à l'aide d'un filtre de Kalman étendu (utilisé dans les travaux de VIANDIER [Viandier 2011]). Notre contribution concerne le développement de deux stratégies d'estimation de la position et leur comparaison à une solution classique de localisation utilisant tous les satellites disponibles. Dans la première solution, nous avons strictement appliqué le principe d'exclusion : tous les satellites considérés comme masqués (dont les signaux ne sont pas reçus en direct) ont été exclus du calcul. Le résultat obtenu en utilisant la stratégie d'exclusion est convaincant lorsque la géométrie des satellites est correcte, mais ne l'est plus lorsque celle-ci est dégradée. Une deuxième stratégie combinant l'information DOP (Dilution of Precision) et l'identification des satellites LOS a été mise en œuvre pour n'appliquer l'exclusion des satellites que lorsque celui-ci est bénéfique. L'information DOP caractérisant la répartition géométrique des satellites, le résultat obtenu en utilisant l'approche de seuillage du DOP est toujours meilleure comparée aux deux autres stratégies (celle basée sur tous les satellites disponibles et celle basée sur l'exclusion des satellites NLOS). La précision des positions estimées est améliorée d'un facteur 2. Les résultats obtenus sont exploités dans les travaux futurs du projet CAPLOC.



## PERSPECTIVES

Les résultats auxquels nous sommes parvenus suscitent des réflexions et ouvrent la voie à plusieurs perspectives qui peuvent faire l'objet de travaux futurs.

En ce qui concerne la combinaison couleur/texture non paramétrique, nous envisageons à court terme, d'enrichir les tests effectués par la mise en œuvre d'autres attributs de texture, notamment les Local Binary Pattern - LBP afin d'améliorer la performance de la combinaison couleur/texture proposée. La combinaison couleur/texture proposée dans cette thèse n'étant effectuée qu'avec des descripteurs calculés dans un espace couleur (nous disposons que de trois valeurs propres), nous envisageons d'utiliser plusieurs espaces couleur à la fois, afin d'avoir un nuage de points à  $N$  dimensions et ainsi une matrice de covariance générique. Nous analyserons ainsi davantage de valeurs propres et le résultat de la combinaison nous informera sur la/les composante(s) couleur la/les plus pertinente(s).

En ce qui concerne l'application visée, il serait intéressant de réaliser une caractérisation plus fine de l'environnement de réception des signaux GNSS. Le passage d'une partition binaire à une division de l'image en trois classes " ciel ", " bâtiments " et " végétation " est donc envisagé afin de pouvoir intégrer/pondérer ces informations dans le calcul d'une nouvelle position. Une caractérisation plus fine des " bâtiments " en utilisant des descripteurs couleur/texture plus évolués pourrait également permettre de déterminer le type des matériaux/surfaces constituant un obstacle. Ceci permettra d'envisager de calculer un indice de réflexion de chaque constituant/surface (verre, béton, etc) de bâtiment afin de mieux pondérer/exploiter les signaux reçus par multi-trajet dans le calcul de la position d'un mobile.

# PUBLICATIONS DE L'AUTEUR

## REVUE INTERNATIONALE (1)

1. Toward accurate localization in guided transport : combining GNSS data and imaging information.  
J. Marais, C. Meurie, D. Attia, Y. Ruichek, A. Flancquart.  
Transportation Research Part C : Emerging Technologies (5-Year IF : 2011 =2.284)  
accepted with minor revisions in April 2013.

## RAPPORTS DE PROJETS ET D'ACTIVITÉS (2)

1. Extraction de descripteurs couleur et texture pour la caractérisation de l'environnement de réception des signaux GNSS  
C. Meurie, D. Attia et Y. Ruichek.  
Livrable 1.2 du projet PREDIT CAPLOC, octobre 2013.
2. Détection de la ligne d'horizon et évaluation du pourcentage de ciel visible par analyse d'images couleur.  
C. Meurie, D. Attia, Y. Ruichek et A. Flancquart.  
Livrable 1.1 du projet PREDIT CAPLOC, octobre 2012.

## COMMUNICATIONS INTERNATIONALES AVEC ACTES ET COMITÉ DE LECTURE (3)

1. Eigen Combination of Colour and Texture Informations for Image Segmentation.  
D. Attia, C. Meurie et Y. Ruichek.  
In Proceedings of International Conference on Image and Signal Processing 2012 (ICISP'2012), Agadir, Morocco.
2. Counting of Satellites with Direct Gnss Signals Using Fisheye Camera : a Comparison of Clustering Algorithms.  
D. Attia, C. Meurie, Y. Ruichek et J. Marais.  
14<sup>th</sup> IEEE Intelligent Transportation Systems Conference 2011 (ITSC'2011), pages 7-12, Washington D.C. USA.
3. Image analysis based real time detection of satellites reception state.  
D. Attia, C. Meurie, Y. Ruichek, J. Marais et A. Flancquart.  
13<sup>th</sup> IEEE Intelligent Transportation Systems Conference 2010 (ITSC'2010), pages 1651-1656, Madère Portugal.

## COMMUNICATIONS NATIONALES AVEC ACTES ET COMITÉ DE LECTURE (3)

1. Classification pixellaire pour la caractérisation de l'état de réception des signaux GNSS dans des images Fisheye.  
D. Attia, C. Meurie, Y. Ruichek et J. Marais.  
XXIII Colloque Traitement du Signal et des Images 2011 (GRETSI'2011), Bordeaux, France.
2. Analyse de l'environnement de réception des signaux GNSS par segmentation et classification d'images Fish-Eye.  
D. Attia, C. Meurie, Y. Ruichek et J. Marais.  
7<sup>ème</sup> Colloque International de TELECOM 2011 (JFMMA'2011), Agadir, Maroc.
3. Une Méthode De Segmentation Hybride Par Combinaison Adaptative Des Informations Texture Et Couleur.  
A. Cohen, D. Attia, C. Meurie et Y. Ruichek. Conference 2010 (MAJESTIC'2010), Bordeaux France.

## COMMUNICATION AVEC ACTES ET SANS COMITÉ DE LECTURE (1)

1. Détection du ciel par analyse d'image Fish-Eye.  
D. Attia.  
Journées des doctorants SPI-STIC de l'INRETS, 2010. ISBN : 978-2-85782-690-3.

# BIBLIOGRAPHIE

- [Adamek 2005] T. Adamek, N. O'Connor et N. Murphy. *Regionbased segmentation of images using syntactic visual features*. In 6th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), 2005.
- [Allili 2007] M. S. Allili et D. Ziou. *Globally adaptive region information for automatic colour–texture image segmentation*. Pattern Recognition Letters, vol. 28, pages 1946–1956, 2007.
- [Angelini 2007] E. Angelini, I. Bloch, Y. Gousseau, S. Ladjal, B. Pesquet-Popescu, M. Sigelle et F. Tupin. *Morphologie mathématique*. In Méthodes Avancées de Traitement d'Images SI 343 - MATIM, chapitre 6, pages 95–126. Polycopié du cours SI343 - MATIM, 2007.
- [Angulo 2003] J. Angulo et J. Serra. *Color segmentation by ordered mergings*. In International Conference on Image Processing, volume 2, pages 125–128, 2003.
- [Angulo 2007] J. Angulo. *Morphological texture gradients. Application to colour+texture watershed segmentation*. 8th International Symposium on Mathematical Morphology, pages 399–410, Octobre 2007.
- [Anys 1995] H. Anys et D.C. He. *Evaluation of textural and multipolarization radar features for crop classification*. IEEE Transactions on Geoscience and Remote Sensing, vol. 23, pages 1169–1181, 1995.
- [Arbelaez 2011] P. Arbelaez, M. Maire, C. Fowlkes et Fellow J. Malik. *Contour Detection and Hierarchical Image Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, pages 898–916, 2011.
- [Bazeille 2004] S. Bazeille, P. Hemar, B. Mougél et F. NG FOK. *Rapport d'images numériques*. Rapport technique, Université de la Rochelle, 2004.
- [Ben-Hur 2002] A. Ben-Hur, A. Elisseeff, A et I. Guyon. *A stability based method for discovering structure in clustered data*. Pacific Symposium on Biocomputing, vol. 7, pages 6–17, 2002.
- [Bennett 1998] J. Bennett et A. Khotanzad. *Multispectral random field models for synthesis and analysis of color images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pages 327–332, 1998.
- [Berkeley] Berkeley. *The Berkeley Segmentation Dataset and Benchmark*. <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>.
- [Bettinger 2012] P. Bettinger et K. L. Merry. *Influence of the juxtaposition of trees on consumer-grade GPS position quality*. In Mathematical and computational forestry & natural-resource sciences, 2012.
- [Beucher 1990] S. Beucher. *Segmentation d'images et morphologie mathématique*. PhD thesis, Ecole des mines de Paris, juin 1990.
- [Beucher 1994] S. Beucher. *Watershed, hierarchical segmentation and waterfall algorithm*. Mathematical morphology and its applications to image and signal processing, pages 69–76, 1994.

- [Beugin 2009] J. Beugin, A. Filip et J. Marais. *Simulation approaches to evaluate dependability of satellite-based positioning services in railway transportation applications*. In ESREL, 2009.
- [Bezdek 1981] J. C. Bezdek. *Pattern recognition with fuzzy objective function algorithms*. New York Plenum Press, 1981.
- [Bihan 2003] N. Le Bihan et S.J. Sangwine. *Quaternion principal component analysis of color images*. IEEE International Conference on Image Processing (ICIP), pages 809–812, 2003.
- [Bolshakova 2003] N. Bolshakova et F. Azuaje. *Cluster validation techniques for genome expression data*. Signal processing, vol. 83, pages 825–833, 2003.
- [Bombardier 2009] V. Bombardier, E. Schmitt et P. Charpentier. *A fuzzy sensor for color matching vision system*. Measurement, vol. 42, pages 189–201, 2009.
- [Borsotti 1998] M. Borsotti, P. Campadelli et R. Shchettini. *Quantitative evaluation of color image segmentation results*. Pattern Recognition Letters, vol. 19, pages 741–747, 1998.
- [Bourdeau 2012] A. Bourdeau, M. Sahmoudi et J-Y. Tournet. *Tight integration of GNSS and a 3D city model for robust positioning in urban canyons*. ION GNSS, 2012.
- [Brox 2010] T. Brox, M. Rousson, R. Deriche et J. Weickert. *Colour, texture and motion in level set based segmentation and tracking*. Image and Vision Computing, vol. 28, pages 376–390, 2010.
- [Bugeau 2011] A. Bugeau. *Détection et suivi d'objets en mouvements dans des scènes complexes, application à la surveillance des conducteurs*. PhD thesis, Université de Rennes 1, 2011.
- [Campbell 1996] N.W. Campbell et B.T. Thomas. *Segmentation of natural images using self organising feature maps*. Proceedings of the Seventh BMVC British Machine Vision Conference, vol. 1, pages 222–232, 1996.
- [Carré 2001] P. Carré, N. Richard et C. Fernandez-Maloigne. *Analyse et caractérisation des signaux bidimensionnels, chapitre Représentation en ondelettes*. Traité I2C, 2001.
- [Carson 2002] C. Carson, S. Belongie, H. Greenspan et J. Malik. *Blobworld : image segmentation using expectation-maximization and its application to image querying*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pages 1026–1038, 2002.
- [Caselles 1997] V. Caselles, R. Kimmel et G. Sapiro. *Geodesic Active Contours*. International Journal of Computer Vision, vol. 22, pages 61–79, 1997.
- [Chabrier 2006] S. Chabrier, B. Emile, C. Rosenberger et H. Laurent. *Unsupervised performance evaluation of image segmentation*. In EURASIP Journal on Applied Signal Processing, 2006.
- [Chan 2001] T. Chan et L. Vese. *Active contours without edges*. In IP, 2001.
- [Chang 2005] Y. Chang, Y. Zhou et Y. Wang. *Combined colour and texture segmentation based on Fibonacci lattice sampling and meanshift*. Proceedings of the Second International Conference on Image Analysis and Recognition (ICIAR), vol. 3656, pages 24–31, 2005.
- [Chau 2007] D. M. Chau et A. Boucher. *Évaluation de la segmentation d'images*. Rapport technique, Institut francophone pour l'informatique, 2007.

- [Chen 2002] K.M. Chen et S.Y. Chen. *Colour texture segmentation using feature distributions*. Pattern Recognition Letters, vol. 23, pages 755–771, 2002.
- [Chen 2004] H. C. Chen et S. J. Wang. *The use of visible color difference in the quantitative evaluation of color image segmentation*. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pages 593–599, 2004.
- [Chen 2005] J. Chen, N. Pappas, Aleksandra Mojsilovic et Bernice E. Rogowitz. *Adaptive Perceptual Color-Texture Image Segmentation*. IEEE Transactions on Image Processing, vol. 14, pages 1–13, 2005.
- [Cheng 2001] H. D. Cheng, X. H. Jiang, Y. Sun et J. Wang. *Color image segmentation : advances and prospects*. Pattern Recognition, vol. 34, pages 2259–2228, 2001.
- [Cheng 2003] J. Cheng, Y.W. Chen, H. Lu et X.Y. Zeng. *Colour and texture based image segmentation using local feature analysis approach*. Proceedings of the SPIE The International Society for Optical Engineering, vol. 5286, pages 600–604, 2003.
- [Chindaro 2006] S. Chindaro, K. Sirlantzis et M.C. Fairhurst. *ICA-based multi-colour space texture classification system*. Electronics Letters, vol. 42(21), pages 1208–1210, 2006.
- [CIE 1998] CIE. *The CIE Interim Colour Appearance Model CIECAM's*. Rapport technique, Rapport technique 131, 1998.
- [CIE 2002] CIE. *A Color Appearance Model for Color Management Systems : CIECAM02*. Rapport technique, Rapport technique, 2002.
- [Cocquerez 1995] J.P. Cocquerez et S. Philipp. *Analyse d'images : filtrage et segmentation*. 1995.
- [Cohen 1989] L. Cohen, L. Vinet, P. Sander et A. Gagalowicz. *Hierarchical regional based stereo matching*. Computer Vision and Pattern Recognition, pages 416–421, 1989.
- [Cohen 2009] A. Cohen, C. Meurie, Y. Ruichek, J. Marais et A. Flanquart. *Quantification of Gns signals accuracy : an image segmentation method for estimating the percentage of sky*. IEEE Internayinal confrenece on Vehicular Electronics and Safety (ICVES), pages 40–45, 2009.
- [Cohen 2010a] A. Cohen, D. Attia, C. Meurie et Y. Ruichek. *Une méthode de segmentation hybride par combinaison adaptative des informations texture et couleur*. In MANifestation des JEunes Chercheurs en Sciences et Technologies de l'Information et de la Communication (MajeSTIC10), 2010.
- [Cohen 2010b] A. Cohen, C. Meurie, Y. Ruichek et J. Marais. *Characterization of the reception environment of GNSS signals using a texture and color based adaptive segmentation technique*. IEEE Intelligent Vehicles Symposium, pages 275–280, june 2010.
- [Comaniciu 1999] D. Comaniciu et P. Meer. *Mean Shift Analysis and Applications*. IEEE International Conference on Computer Vision (ICCV), Kerkyra, Greece, pages 1197–1203, 1999.
- [Comaniciu 2001] D. Comaniciu, V. Ramesh et P. Meer. *The Variable Bandwidth Mean Shift and Data-Driven Scale Selection*. IEEE Int. Conf. Computer Vision (ICCV), Vancouver, Canada, vol. 1, pages 438–445, 2001.



- [Comaniciu 2002] D. Comaniciu et P. Meer. *Mean shift : A robust approach toward feature space analysis*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pages 603–619, 2002.
- [Correia 2003] P. Correia et F. Pereira. *Objective evaluation of video segmentation quality*. IEEE Transactions on Image Processing, vol. 12, pages 186–200, 2003.
- [Correia 2006] P. Correia et F. Pereira. *Video object relevance metrics for overall segmentation quality evaluation*. EURASIP Journal on Applied Signal Processing, 2006.
- [Cour 2005] T. Cour, F. Benezit et J. Shi. *Spectral segmentation with multiscale graph decomposition*. In CVPR, 2005.
- [Cremers 2007] D. Cremers, M. Rousson et R. Deriche. *A review of statistical approaches to level set segmentation : integrating colour, texture, motion and shape*. International Journal of Computer Vision, vol. 72, pages 195–215, 2007.
- [Dash 2012] A. Dash, P. Kanungo et B.P. Mohanty. *A Modified Gray level Co-occurrence Matrix based Thresholding for Object Background Classification*. International Conference on Communication Technology and System Design 2011, Procedia Engineering, vol. 30, pages 85–91, 2012.
- [Datar 2008] M. Datar, D. Padfield et H. Cline. *Colour and texture based segmentation of molecular pathology images using HSOMS*. IEEE International Symposium on Biomedical Imaging : From Macro to Nano (ISBI), pages 292–294, 2008.
- [Deng 2001] Y. Deng et B.S. Manjunath. *Unsupervised Segmentation of Color-Texture Regions in Images and Video*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, pages 800–810, 2001.
- [Destrempe 2006] F. Destrempe, J.F. Angers et M. Mignotte. *Fusion of hidden Markov random field models and its Bayesian estimation*. IEEE Transactions on Image Processing, vol. 15, pages 2920–2935, 2006.
- [DiZenzo 1986] S. DiZenzo. *A note on the gradient of multi-image*. Computer Vision, Graphics, and Image Processing, vol. 33, pages 116–125, 1986.
- [Dongen 2000] S. Van Dongen. *Performance criteria for graph clustering and markov cluster experiments*. Rapport technique, INS-R0012, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, The Netherlands, 2000.
- [Donoser 2009] M. Donoser, M. Urschler, M. Hirzer et H. Bischof. *Saliency driven total variational segmentation*. In ICCV, 2009.
- [Du 2009] Q. Du, W. Zhu, H. Yang et J-E. Fowler. *Segmented Principal Component Analysis for Parallel Compression of Hyperspectral Imagery*. IEEE Geoscience and Remote Sensing Letters, vol. 6, pages 713–717, 2009.
- [Dubois 2008] J. Dubois. *Segmentation par approche contours*. Master's thesis, Université de Rouen, 2008.
- [Dubry 2006] C. Dubry et S. Robin. *Analyse en composantes principales*. Rapport technique, Institut National Agronomique Paris - Grignon, 2006. Département O.M.I.P, Institut National Agronomique Paris - Grignon.
- [Dunn 1973] J. C. Dunn. *A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters*. Journal of Cybernetics, vol. 3, pages 32–52, 1973.

- [Echigo 2000] T. Echigo et S. Lisaku. *Unsupervised segmentation of coloured texture images by using multiple GMRF models and a hypothesis of merging primitives*. Systems and Computers in Japan, vol. 31, pages 29–39, 2000.
- [Everingham 2009] M. Everingham, L. VanGool, C. K. I. Williams, J. Winn et A. Zisserman. *The PASCAL Visual Object Classes Challenge 2009 (VOC2009) Results*. <http://www.pascal-network.org/challenges/VOC/voc2009/workshop/index.html>, 2009.
- [Faugeras 1979] O. D. Faugeras. *Digital color image processing within the framework of a human visual model*. IEEE ASSP, vol. 27, pages 380–393, 1979.
- [Felzenszwalb 2004] P. Felzenszwalb et D. Huttenlocher. *Efficient graph-based image segmentation*. In IJCV, 2004.
- [Fisher 1936] R. Fisher. *The use of multiple measurements in taxonomic problems*. Anals of Eugenics, vol. 7, pages 179–188, 1936.
- [Fisher 1958] W. Fisher. *On grouping for maximum homogeneity*. JASA, vol. 53, pages 789–798, 1958.
- [Fondon 2006] I. Fondon, C. Serrano et B. Acha. *Colour-texture image segmentation based on multistep region growing*. Optical Engineering, vol. 45, 2006.
- [Fortune 1987] S. Fortune. *A Sweepline Algorithm for Voronoi Diagrams*. Algorithmica, vol. 2, pages 153–174, 1987.
- [Fowlkes 1983] E. Fowlkes et C. Mallows. *A method for comparing two hierarchical clusterings*. Journal of the American Statistical Association, vol. 78, pages 553–569, 1983.
- [Fredembach 2004] C. Fredembach, M. Shroder et S. Susstrunk. *Eigneregions for Image Classification*. IEEE Tansaction on Pattern Analysis and Machine Intelligence, vol. 26, pages 1645–1649, 2004.
- [Freixenet 2004] J. Freixenet, X. Munoz, J. Marti et X. Llado. *Colour texture segmentation by region-boundary cooperation*. Proceedings of the European Conference on Computer Vision, vol. 2, pages 250–261, 2004.
- [Fu 1981] K. Fu et J. Mui. *A survey on image segmentation*. Pattern Recognition, vol. 13, pages 3–16, 1981.
- [Fukuda 2009] K. Fukuda, T. Takiguchi et Y. Ariki. *Graphcuts segmentation by using local texture features of multiresolution analysis*. IEICE Transactions on Information and Systems, vol. E92-D, pages 1453–1461, 2009.
- [Fukunaga 1975] K.. Fukunaga et L.D. Hostler. *Estimation of the gradient of a density function with applications in pattern recognition*. IEEE Trancactions on Information Theory, vol. 21, pages 32–40, 1975.
- [Galloway 1975] M-M. Galloway. *Texture analysis using gray level run lengths*. Computer graphics and image processing, vol. 4, pages 172–179, 1975.
- [Garcia 2008] R. L. Garcia, R. Deriche et C. Alberola-Lopez. *Texture and colour segmentation based on the combined use of the structure tensor and the image components*. Signal Processing, vol. 88, pages 776–795, 2008.
- [Ge 2007] F. Ge, S. Wang et T. Liu. *New benchmark for image segmentation evaluation*. Journal of Electronic Imaging, vol. 16, pages 033011–1–033011–16, 2007.
- [Gergaud 2006] J. Gergaud. *Unité Fondamentale : Algèbre linéaire : une application l'Analyse en Composantes Principales*. Rapport technique, INP ENSAT, 2006.

- [Gevers 2002] T. Gevers. *Image segmentation and similarity of colour–texture objects*. IEEE Transactions on Multimedia, vol. 4, pages 509–516, 2002.
- [Griniasa 2008] I. Griniasa, N. Komodakis et G. Tziritas. *Bayesian region growing and MRF-based minimisation for texture and colour segmentation*. 8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), page 20, 2008.
- [Groves 2011] P. D. Groves. *Shadow matching : a new GNSS positioning technique for urban canyons*. The journal of Navigation, vol. 64, pages 417–430, 2011.
- [Guigue 2003] V. Guigue, A. Rakotomamonjy et S. Canu. *SVM et k-ppv pour la reconnaissance d'émotions*. In 19 ème Colloque sur le traitement du signal et des images, pages 770–773. Groupe d'Etudes du Traitement du Signal et des Images GRETSI, 2003.
- [Gupta 2008] L. Gupta, U. Gosa M. et S. Das. *Integrating region and edge information for texture segmentation using a modified constraint satisfaction neuralnetwork*. Image and Vision Computing, vol. 26, pages 1106–1117, 2008.
- [Haindl 2008] M. Haindl et S. Mikes. *Texture segmentation benchmark*. In 19 th International Conference on Pattern Recognition, pages 1–4, 2008.
- [Hamming 1950] R. Hamming. *Error-detecting and error-correcting codes*. Bell System Technical Journal, vol. 29, pages 147–160, 1950.
- [Han 2009] S. Han, W. Tao, D. Wang, X.C. Tai et X. Wu. *Image segmentation based on GrabCut framework integrating multiscale non linear structure tensor*. IEEE Transactions on Image Processing, vol. 18, pages 2289–2302, 2009.
- [Hanbury 2002] A. Hanbury. *Morphologie mathématique sur le cercle unité avec applications aux teintes et aux textures orientées*. PhD thesis, École supérieure des Mines de Paris, 2002.
- [Hanbury 2009] A. Hanbury et B. Marcotegui. *Morphological segmentation on learned boundaries*. Image and Vision Computing, vol. 27, pages 480–488, 2009.
- [Haralick 1973] R. M. Haralick, K. Shanmugan et I. Dinstein. *Textural features for image classification*. IEEE Trans. On Systems, Man, and Cybernetics, vol. 3, no. 6, pages 610–621, Novembre 1973.
- [Haralick 1979] R. M. Haralick. *Statistical and Structural Approaches to Texture*. Proceedings of the IEEE, vol. 67, pages 786–804, 1979.
- [Harwood 1993] D. Harwood, T. Ojala, M. Pietikainen, S. Kelman et S. Davis. *Texture classification by center symmetric auto correlation using Kullback discrimination of distributions*. Rapport technique, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, Maryland. CAR-TR-678, 1993.
- [Hays 2008] J. Hays et A. A. Efros. *Im2gps : estimating geographic images from single images*. In Computer Vision and Pattern Recognition (CVPR), 2008.
- [Hedjam 2009] R. Hedjam et M. Mignotte. *Ahierarchical graph-based Markovian clustering approach for the unsupervised segmentation of textured colour images*. Proceedings of the International Conference on Image Processing (ICIP), pages 1365–1368, 2009.
- [Hering 1880] E. Hering. *Zur Erklärung der Farbenblindheit aus der Theorie der Gegenfarben*. Lotos, Jb. f. Naturwiss, vol. 1, pages 76–107, 1880.

- [Herlidou 1999] S. Herlidou. *Caractérisation tissulaire en IRM par l'analyse de texture : Etude du tissu musculaire et de tumeurs intracrâniennes*. PhD thesis, Université de Rennes 1, 1999.
- [Hernandez 2004] O.J. Hernandez et A. Khotanzad. *Color image segmentation using multispectral random field texture model and color content features*. Journal of Computer Science et Technology, vol. 4, pages 141–146, 2004.
- [Hernandez 2005] O.J. Hernandez, J. Cook, M. Griffin, C. De Rama et M. McGovern. *Classification of color textures with random field models and neural networks*. Journal of Computer Science et Technology, vol. 5, pages 150–157, 2005.
- [Hibon 2010] O. Hibon. *Analyse d'image : caractérisation de textures d'images de gels lactés*. Master's thesis, Institut National de Recherche Agronomique, 2010.
- [Ho 2005] Pun-Mo Ho, Tien-Tsin Wong et Chi-Sing Leung. *Compressing the Illumination-Adjustable Images With Principal Component Analysis*. IEEE Transaction on Circuits and Systems for Video Technology, vol. 15, pages 355–364, 2005.
- [Hoang 2005] M.A. Hoang, J.M. Geusebroek et A.W. Smeulders. *Colour texture measurement and segmentation*. Signal Processing, vol. 85, pages 265–275, 2005.
- [Huang 1992] C.L. Huang, T.Y. Cheng et C.C. Chen. *Colour images segmentation using scale space filter and Markov random field* 1217–1229. Pattern Recognition, vol. 25, pages 1217–1229, 1992.
- [Huang 1995] Q. Huang et B. Dom. *Quantitative methods of evaluating image segmentation*. In International Conference on Image Processing (ICIP), pages 53–56, 1995.
- [Huawu 2004] D. Huawu et D.A. Clausi. *Unsupervised image segmentation using a simple MRF model with an ew implementation scheme*. Pattern Recognition, vol. 37, pages 2323–2335, 2004.
- [Hurvich 1957] L. M. Hurvich et D. Jameson. *An opponent-process theory of color vision*. Psychological Review, vol. 64, pages 384–404, 1957.
- [Ilea 2006] D.E. Ilea et P.F. Whelan. *Colour image segmentation using a self-initializing EM algorithm*. In Proceedings of the International Conference on Visualisation Imaging and Image Processing (VIIP), 2006.
- [Ilea 2007] D.E. Ilea et P.F. Whelan. *Adaptive pre-filtering techniques for colour image analysis*. IEEE Computer Society Press Proceedings of the International Machine Vision and Image Processing Conference (IMVIP2007), pages 150–157, 2007.
- [Ilea 2008] D.E. Ilea et P.F. Whelan. *CTex—an adaptive unsupervised segmentation algorithm based on colour–texture coherence*. IEEE Transaction on Image Processing, vol. 17, pages 1926–1939, 2008.
- [Ilea 2009] D.E. Ilea et P.F. Whelan. *Colour saliency-based parameter optimisation for adaptive colour segmentation*. IEEE International Conference on Image Processing, pages 973–976, 2009.
- [Ilea 2011] D. E. Ilea et P. F. Whelan. *Image segmentation based on the integration of colour-texture descriptors*. Pattern Recognition, vol. 44, pages 2479–2501, 2011.
- [Ilin 2010] A. Ilin et T. Raiko. *Practical Approaches to Principal Component Analysis in the Presence of Missing Values*. In Journal of Machine Learning Research 11, 2010.

- [Jacobs 2007] N. Jacobs, S. Satkin, N. Roman, R. Speyer et R. Pless. *Geolocating static cameras*. In International Conference on Computer Vision (ICCV), 2007.
- [Jain 1998] A.K. Jain et R.C. Dubes. *Algorithms for clustering data*. Prentice-Hall, 1998.
- [Jiang 2006] X. Jiang, C. Marti, C. Irniger et H. Bunke . . *Distance measures for image segmentation evaluation*. In EURASIP Journal on Applied Signal Processing, 2006.
- [Jolly 2000] M.P.D. Jolly et A.Gupta. *Colour and texture fusion : application to aerial image segmentation and GIS updating*. Image and Vision Computing, vol. 18, pages 823–832, 2000.
- [Julesz 1981] B. Julesz. *Textons, the elements of texture perception and their interactions*. Nature, vol. 290, pages 91–97, 1981.
- [Kailath 1967] T. Kailath. *The Divergence and Bhattacharyya Distance Measures in Signal Selection*. IEEE Transactions on Communication Technology, vol. 15, pages 52–60, 1967.
- [Kanungo 2003] T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman et A. Y. Wu. *A Local Search Approximation Algorithm for k-Means Clustering*. Computational Geometry : Theory and Applications, vol. 28, pages 89–112, 2003.
- [Kato 2002] Z. Kato, T.C. Pong et S.G. Qiang. *Multicue MRF image segmentation : combining texture and colour features*. Proceedings of the 16 th International Conference on Pattern Recognition, vol. 1, pages 660–663, 2002.
- [Kato 2006] Z. Kato et T.C. Pong. *A Markov random field image segmentation model for colour textured images*. Image and Vision Computing, vol. 24, pages 1103–1114, 2006.
- [Khan 2009] J.F. Khan, R.R. Adhami et S.M.A. Bhuiyan. *A customized Gabor filter for unsupervised colour image segmentation*. Image and Vision Computing, vol. 27, pages 489–501, 2009.
- [Khotanzad 2003] A. Khotanzad et O.J. Hernandez. *Colour image retrieval using multispectral random field texture model and colour content features*. Pattern Recognition, vol. 36, pages 1679–1694, 2003.
- [Khotanzad 2006] A. Khotanzad et O.J. Hernandez. *A classification methodology for color textures using multispectral random field mathematical models*. Mathematical and Computational Applications, vol. 11, pages 111–120, 2006.
- [Kim 2009] J.S. Kim et K.S.C Hong. *Colour–texture segmentation using unsupervised graph cuts*. Pattern Recognition, vol. 42, pages 735–750, 2009.
- [Koch 2007] O. Koch et S. Teller. *Wide-area egomotion estimation from known 3d structure*. In Conference on Computer Vision and Pattern Recognition (CVPR), 2007.
- [Krinidis 2009] M. Krinidis et I. Pitas. *Colour texture segmentation based on the modal energy of deformable surfaces*. IEEE Transactions on Image Processing, vol. 18, pages 1613–1622, 2009.
- [Kullback 1951] S. Kullback et R. Leibler. *On information and sufficiency*. Annals of Mathematical Statistics, vol. 22, pages 79–86, 1951.
- [Kyperountasa 2008] M. Kyperountasa, A. Tefasa et I. Pitasa. *Dynamic training using multistage clustering for face recognition*. Pattern Recognition, vol. 41, pages 894–905, 2008.

- [Lafarge 2004] F. Lafarge, X. Descombes et J. Zerubia. *Noyaux texturaux pour les problèmes de classification par SVM en télédétection*. Rapport technique 5370, INRIA, novembre 2004.
- [Lambert 2002] P. Lambert. *Etudes méthodologiques du filtrage et de la segmentation d'images multi-composantes*. PhD thesis, UNIVERSITE DE SAVOIE, juillet 2002. Habilitation a Diriger Des Recherches en Electronique, Electrotechnique et Automatique.
- [Lefèvre 2009] S. Lefèvre. *A New Approach for Unsupervised Classification in Image Segmentation*. EGC, pages 127–138, 2009.
- [Lentmaier 2008] M. Lentmaier, B. Krach et P. Robertson. *Bayesian Time delay estimation of GNSS signals in dynamic multipath environments*. International Journal of Navigation and Observation, vol. 2008, no. ID 372651, pages 1–12, 2008.
- [Levine 1985] M. D. Levine et A. M. Nazif. *Dynamic measurement of computer generated image segmentations*. PAMI, vol. 7, pages 155–164, 1985.
- [Lezoray 2008] O. Lezoray, C. Meurie et A. Elmoataz. *Graph-based ordering scheme for color image filtering*. International Journal of Image and Graphics, vol. 8, no. 3, pages 473–493, Juillet 2008.
- [Liu 1994] J. Liu et Y. H. Yang. *Multiresolution color image segmentation*. IEEE Transactions on PAMI, vol. 16, pages 689–700, 1994.
- [Lloyd 1982] S. P. Lloyd. *Least squares quantization in PCM*. IEEE Transactions on Information Theory, vol. 28, pages 129–137, 1982.
- [Lo 2012] J. T-H Lo. *A cortex-like learning machine for temporal hierarchical patternclustering, detection, and recognition*. Neurocomputing, vol. 78, pages 89–103, 2012.
- [Lughofer 2008] E. Lughofer. *Extensions of vectorquantization for incremental clustering*. Pattern Recognition, vol. 41, pages 995–1011, 2008.
- [Maciejewski 2001] A. A. Maciejewski et R. G. Roberts. *An Example of Principal Component Analysis Applied to Correlated Images*. In IEEE Proceedings of the 33rd Southeastern Symposium, 2001.
- [MacQueen 1967] J. MacQueen. *Some methods for classification and analysis of multivariate observations*. In Proceedings of the Fifth Berkley Symposium on Mathematical Statistics and Probability, vol. 1, pages 281–296, 1967.
- [Majdoulayne 2009] H. Majdoulayne. *Extraction de caractéristiques de texture pour la classification d'images satellites*. PhD thesis, Université de Toulouse, 2009.
- [Manduchi 1999] R. Manduchi. *Bayesian fusion of colour and texture segmentations*. Proceedings of the IEEE International Conference on Computer Vision, vol. 2, pages 956–962, 1999.
- [Manzanera ] A. Manzanera. *Traitement d'images et vision artificielle*. Unité d'Électronique et d'Informatique.
- [Marais 2002] J. Marais, M. Berbineau, J-P. Ghys et M. Heddebaut. *Effect of multipath on availability of GPS satellites*. In International Communications Satellite Systems Conference (ICSSC 2002), Montréal, 2002.
- [Marais 2012] J. Marais, S. Ambellouis, A. Flancquart, S; Lefebvre et Y. Ruichek C. Meurie. *Accurate Localisation Based on GNSS and Propagation Knowledge for Safe Applications in Guided Transport*. Procedia - Social and Behavioral Sciences, vol. 48, pages 796–805, 2012.



- [Martin 2001] D. Martin, C. Fowlkes, D. Tal et J. Malik. *A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics*. In Proc. 8th Int'l Conf. Computer Vision, 2001.
- [Martin 2002] D. Martin. *An Empirical Approach to Grouping and Segmentation*. PhD thesis, University of California Berkeley, 2002.
- [Martin 2004] D.R. Martin, C.C. Fowlkes et J. Malik. *Learning to detect natural image boundaries using local brightness, colour, and texture cues*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pages 530–549, 2004.
- [Mathankera 2011] S.K. Mathankera, P.R. Wecklera, T.J. Bowsera, N. Wanga et N.O. Manessb. *AdaBoost classifiers for pecan defect classification*. Computers and Electronics in Agriculture, vol. 77, pages 60–68, 2011.
- [Mavromatis 2001] S. Mavromatis. *Analyse de texture et visualisation scientifique*. PhD thesis, Université de la méditerranée, 2001.
- [McGuinness 2009] K. McGuinness. *Image Segmentation, Evaluation, and Applications*. PhD thesis, School of Electronic Engineering, 2009.
- [Meguro 2008] J-I Meguro, T. Murata, J-I Takiguchi, Y. Amano et T. Hashizume. *GPS Accuracy Improvement by Satellite Selection Using Omnidirectional Infrared Camera*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 22–26, Septembre 2008.
- [Meguro 2009] J-I Meguro, Taishi Murata, J-I Takiguchi, Yoshiharu Amano et Takumi Hashizume. *GPS Multipath Mitigation for Urban Area Using Omnidirectional Infrared Camera*. IEEE Trans. On Intelligent Transportation Systems, vol. 10, no. 1, pages 22–30, Mars 2009.
- [Meila 2003] M. Meila. *Comparing clusterings by the variation of information*. The 16th Annual Conference on Computational Learning Theory and 7th Workshop on Kernel Machines (COLT/Kernel), pages 173–187, 2003.
- [Meurie 2005] C. Meurie. *Segmentation of color images by pixels classification and hierarchy of partitions*. PhD thesis, University of Caen Basse-Normandie, Caen, France, Octobre 2005.
- [Meurie 2010a] C. Meurie, A. Cohen et Y. Ruichek. *An Efficient Combination of Texture and Color Information for Watershed Segmentation*. In Proceedings of International Conference on Image and Signal Processing ICISP, pages 147–156, 2010.
- [Meurie 2010b] C. Meurie, Y. Ruichek, A. Cohen et J. Marais. *An hybrid an adaptive segmentation method using color and textural information*. In Image Processing : Machine Vision Applications III, Janvier 2010.
- [Mezaris 2003] V. Mezaris, I. Kompatsiaris et M. G. Strintzis. *Still image objective segmentation evaluation using ground truth*. Fifth COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication, pages 9–14, 2003.
- [Mignotte 2008] M. Mignotte. *Segmentation by Fusion of Histogram-Based K-Means Clusters in Different Color Spaces*. IEEE Transactions On Image Processing, vol. 17, pages 780–787, 2008.
- [Mignotte 2010] M. Mignotte. *A label field fusion Bayesian model and its penalized maximum Rand estimator for image segmentation*. IEEE Transactions on Image Processing, vol. 19, pages 1610–1624, 2010.

- [Mirmehdi 2000] M. Mirmehdi et M. Petrou. *Segmentation of colour textures*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pages 142–159, 2000.
- [Moosmann 2008] F. Moosmann, E. Nowak et F. Jurie. *Randomized Clustering Forests for Image Classification*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pages 1632–1646, 2008.
- [Morris 1986] O. Morris, M. Lee et A. Constantinides. *Graph theory for image analysis : an approach based on the shortest spanning tree*. IEE Proceedings F. Communications, Radar and Signal Processing, vol. 133, pages 146–152, 1986.
- [Mount 2005] D. M. Mount. *KMlocal : A Testbed for k-means Clustering Algorithms*. Rapport technique, Department of Computer Science and Institute for Advanced Computer Studies University of Maryland, 2005.
- [Moussafir 2005] J. O. Moussafir. *Introduction aux SVM*. 2005. 20 octobre 2005.
- [Mudrova ] M. Mudrova et A. Prochazka. Principal component analysis in image processing. Departement of Computing and Control Engineering, Institute of Chemical Technology, Prague.
- [Mumford 1989] D. Mumford et J. Shah. *Optimal approximations by piecewise smooth functions, and associated variational problems,*. Communications on Pure and Applied Mathematics, pages 577–684, 1989.
- [Najman 2008] L. Najman et H. Talbot. Morphologie mathématique : Tome 1, approches déterministes, volume 1 of *IC2 : Série Traitement du signal et de l'image*. Hermès-Lavoisier, 1 édition, 10 2008.
- [Nammalwar 2004] P. Nammalwar, O. Ghita et P.F. Whelan. *Integration of feature distributions for colour texture segmentation*. Proceedings of the 17 th International Conference on Pattern Recognition (ICPR), vol. 1, pages 716–719, 2004.
- [Nammalwar 2010] P. Nammalwar, O. Ghita et P.F. Whelan. *A generic framework for colour texture segmentation*. Sensor Review, vol. 30, pages 69–79, 2010.
- [Niea 2011] F. Niea, C. Gaoa, Y. Guoa et M. Ganc. *Two-dimensional minimum local cross-entropy thresholding based on co-occurrence matrix*. Computers and Electrical Engineering, vol. 37, pages 757–767, 2011.
- [Nielsen 2003] F. Nielsen et R. Nock. *On region merging : the statistical soundness of fast sorting, with applications*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pages 19–26, 2003.
- [Niskanen 2001] M. Niskanen, O. Silven et H. Kauppinen. *Colour and texture based wood inspection with non-supervised clustering*. in : Proceedings of the 12th Scandinavian Conference on Image Analysis (SCIA01),, pages 336–342, 2001.
- [Nock 2004a] R. Nock et F. Nielsen. *Statistical Region Merging*. IEEE transaction on pattern analysis and machine intelligence, vol. 26, no. 11, pages 1452–1458, november 2004.
- [Nock 2004b] R. Nielsen F. Nock. *Grouping with bias revisited*. Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 2, pages 460–465, 2004.
- [Nock 2005] R. Nock et F. Nielson. *Semi-supervised statistical region refinement for color image segmentation*. Pattern Recognition, vol. 38, pages 835–846, 2005.

- [Nowak 2005] D. Nowak. Calibrage d'images issues d'une caméra équipée d'un objectif fisheye. Master's thesis, INRETS, juillet 2005.
- [Obst 2012] M. Obst, S. Bauer et G. Wanielik. *Urban multipath detection and mitigation with dynamic 3D maps for reliable land vehicle localization*. PLANS, 2012.
- [Ojala 1996] T. Ojala, M. Pietikinen et D. Harwood. *A comparative study of texture measures with classification based on feature distributions*. Pattern Recognition, vol. 29, pages 51–59, 1996.
- [Ojala 2002] T. Ojala, T. Maenpää, M. Pietikainen, J. Viertola, J. Kyllönen et S. Huovinen. *Outex—new framework for empirical evaluation of texture analysis algorithms*. 16th International Conference on Pattern Recognition (ICPR) Quebec, Canada, vol. 1, pages 701–706, 2002.
- [Olmos 2004] A. Olmos et F. Kingdom. McGill calibrated colour image database. 2004.
- [Ooi 2009] W.S. Ooi et C.P. Lim. *Fusion of colour and texture features in image segmentation : an empirical study*. Imaging Science Journal, vol. 57, pages 8–18, 2009.
- [Osei-Bryson 2010] K-M. Osei-Bryson. *Towards supporting expert of clustering results using a data mining process model*. Information Sciences, vol. 180, pages 414–431, 2010.
- [Ozyildiz 2002] E. Ozyildiz, N. Krahnstover et R. Sharma. *Adaptive texture and colour segmentation for tracking moving objects*. Pattern Recognition, vol. 35, pages 2013–2029, 2002.
- [Palm 2000] C. Palm, D. Keysers, T. Lehmann et K. Spitzer. *Gabor filtering of complex hue saturation images for color texture classification*. Proceedings of the 5th Joint Conference and Information Science, pages 45–49, 2000.
- [Palm 2002] C. Palm et T. Lehmann. *Classification of color textures by gabor filtering*. Machine Graphics & Vision International Journal, vol. 11, pages 195–219, 2002.
- [Palm 2004] C. Palm. *Color texture classification by integrative co-occurrence matrices*. Pattern recognition, vol. 37, pages 965–976, 2004.
- [Panjwani 1993] D.K. Panjwani et G. Healey. *Unsupervised segmentation of textured colour images using Markov random field models*. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 776–777, 1993.
- [Pappas 1992] T.N. Pappas. *An adaptive clustering algorithm for image segmentation*. IEEE Transactions on Image Processing, vol. 14, pages 901–914, 1992.
- [Paquis 2003] S. Paquis. *Reconnaissance des surfaces de chaussée par analyse de texture image*. PhD thesis, Université Jean Monnet de Saint-Étienne, 2003.
- [Paschos 1995] G. Paschos et K.P. Valavanis. *Chromatic measures for colour texture description and analysis*. Proceedings of the IEEE International Symposium on Intelligent Control, pages 319–325, 1995.
- [Paschos 1999] G. Paschos et F.P. Valavanis. *A colour texture based visual monitoring system for automated surveillance*. IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews), vol. 29, pages 298–307, 1999.
- [Perona 1990] M. Perona et J. Malik. *Scale-space and edge detection using anisotropic diffusion*. IEEE Trans. PAMI, vol. 12, pages 629–639, 1990.

- [Peyret 2011] F. Peyret, D. Bétaille et F. Mougel. *Non-Line-Of-Sight GNSS signal detection using an on-board 3D model of buildings*. In ITST conference, St-Petersburg, 2011.
- [Philpp-Foliguet 2006] S. Philpp-Foliguet et L. Guiges. *Évaluation de la segmentation d'images : état de l'art, nouveaux indices et comparaison*. GRETSI, Saint Martin d'Hères, France, vol. 23, pages 109–124, 2006.
- [Pietikäinen 2002] M. Pietikäinen, T. Mäenpää et J. Viertola. *Color texture classification with color histograms and local binary patterns*. In Proceedings of the 2nd International Workshop on Texture Analysis and Synthesis, pages 109–112, 2002.
- [Pietikäinen 2011] M. Pietikäinen, A. Hadid, G. Zhao et T. Ahonen. Computer vision using local binary patterns, volume 40, chapitre Local Binary Patterns for Still Images, pages 13–47. Springer-Verlag London Limited, 2011.
- [Pire ] B. Pire. *Experience de YOUNG*. Encyclopædia Universalis.
- [Porebski 2007] A. Porebski et N. Vandenbroucke L. Macaire. *Iterative Feature Selection for Color Texture Classification*. IEEE International Conference on Image Processing ICIP, vol. 3, pages 509–512, 2007.
- [Porebski 2009] A. Porebski, N. Vandenbroucke et L. Macaire. *Attributs de texture couleur*. In Proceedings de l'École d'Hiver sur l'Imagerie Numérique Couleur (EHINC'09), 2009. Ecole d'Ingénieurs du Pas-de-Calais (EIPC), Laboratoire d'Automatique, G'éenie Informatique et Signal (LAGIS UMR CNRS ), Univeristé des Sciences et Technologies de Lille.
- [Pratt 1978] W. Pratt. Digital image processing. New York, 1978.
- [ProFlex 2009] ProFlex. *Reference Manual*. Rapport technique, MAGELLAN Professional, 2009.
- [Qazi 2009] I.U.H. Qazi, O. Alata, J.C. Burie et C. Fernandez-Maloigne. *Colour spectral analysis for spatial structure characterization of textures in IHLS colour space*. Pattern Recognition, vol. 43, pages 663–675, 2009.
- [Ramalingam 2010] S. Ramalingam, S. Bouaziz, P. Sturm et M. Brand 1. *SKYLINE2GPS : Localization in Urban Canyons using Omni-Skylines*. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010.
- [Rand 1971] W.M. Rand. *Objective criteria for the evaluation of clustering methods*. Journal of the American Statistical Association, vol. 66, pages 846–850, 1971.
- [Rao 2009] S. Rao, H. Mobahi, A. Yang, S. Sastry et Y. Ma. *Natural Image Segmentation with Adaptive Texture and Boundary Encoding*. In ACCV, 2009.
- [Rivest 1993] J. F. Rivest, P. Soille et S. Beucher. *Morphological gradients*. Journal of. Electronic Imaging, vol. 2, pages 326–336, 1993. Centre de Morphologie Mathématique, Ecole des Mines de Paris.
- [Robertson 2004] D. Robertson et R. Cipolla. *An image-based system for urban navigation*. In British Machine Vision Conference (BMVC), 2004.
- [Rogowitz 2005] J. Chen T.N. Pappas A. Mojsilovic B.E. Rogowitz. *Adaptive perceptual colour–texture image segmentation*. IEEE Transactions on Image Processing, vol. 14, pages 1524–1536, 2005.
- [Royer 2007] E. Royer, M. Lhuillier et M. Dhome. *Monocular vision for mobile robot localization*. In IJCV, 2007.

- [Savitzky 1964] A. Savitzky et M. J. E. Golay. *Smoothing and differentiation of data by simplified least squares procedures*. Analytical Chemistry, vol. 36, pages 1627–1639, 1964.
- [Sbeh 2001] Z. B. Sbeh et L. D. Cohen. *A new approach of geodesic reconstruction for Drusen segmentation in eye funds images*. IEEE Transactions on medical imaging, vol. 20, pages 1321–1333, 2001.
- [Scarpa 2006] G. Scarpa et M. Haindl. *Unsupervised texture segmentation by spectral-spatial-independent clustering*. Proceedings of the 18th International Conference on Pattern Recognition (ICPR), vol. 2, pages 151–154, 2006.
- [Schmitt 1994] M. Schmitt et J. Mattioli. Morphologie mathématique. Paris, 1994.
- [Schubert 2009] F. M. Schubert, A. Lehner, A. Steingass, P. Robertson, B. H. Fleury et R. Prieto-Cerdeira. *Modeling the GNSS Rural Radio Channel : Wave Propagation Effects caused by Trees and Alleys*. In ION 2009 Savannah, GA, 2009.
- [Schubert 2010] F.M. Schubert, B. H. Fleury, P. Robertson, R. Prieto-Cerdeira, A. Steingass et A. Lehner. *Modeling of multipath propagation components caused by trees and forests*. In Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP), pages 1–5, april 2010.
- [Schubert 2012] F. M. Schubert, R. Prieto-Cerdeira B. H. Fleury, A. Steingass et A. Lehner. *A Rural Channel Model for Satellite Navigation Applications*. In EuCAP 2012, Prague, march 2012.
- [Selim 1984] S.Z. Selim et M. A. Ismail. *K-means-type algorithms : A generalized convergence theorem and characterization of local optimality*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, pages 81–87, 1984.
- [Serra 1988] J. Serra. *Image analysis and mathematical morphology*. Theoretical advances, vol. 2, 1988.
- [Serrano 2009] C. Serrano et B. Acha. *Pattern analysis of dermoscopic images based on Markov random fields*. Pattern Recognition, vol. 42, pages 1052–1057, 2009.
- [Shafarenko 1997] L. Shafarenko, M. Petrou et J. Kittler. *Automatic watershed segmentation of randomly textured color images*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 6, no. 11, pages 1530–1544, november 1997.
- [Sharma 2001a] M. Sharma. *Evaluation of image segmentation and texture extraction methods in scene analysis*. PhD thesis, University of Exeter, 2001.
- [Sharma 2001b] M. Sharma et S. Singh. *Evaluation of texture methods for image analysis*. In 7th Australian and New Zealand Intelligent Information Systems Conference, pages 117–121, 2001.
- [Shi 2000] J. Shi et J. Malik. *Normalized cuts and image segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pages 888–905, 2000.
- [Shi 2005] L. Shi et B. Funt. *Quaternion colour texture*. In 10th Congress of the International Color Association,, 2005.
- [Shi 2007] L. Shi et B. Funt. *Quaternion color texture segmentation*. Computer Vision and Image Understanding, vol. 107, pages 88–96, 2007.
- [Shmitt 1994] M. Shmitt et J. Mattioli. Morphologie mathématique. Masson, 1994.
- [Siroyich 1987] L. Siroyich et M. Kirby. *Low-dimensional procedure for the characterization of human faces*. Journal of Optical Society of America, vol. 4, pages 519–524, 1987.

- [Skarbek 1994] W. Skarbek et A. Koschan. *Colour Image Segmentation : A Survey*, 1994.
- [Sklansky 1978] J. Sklansky. *Image segmentation and feature extraction*. IEEE Transactions on Systems, Man, and Cybernetics, vol. 8, pages 237–247, 1978.
- [Sofou 2005] A. Sofou et G. Evan. *Coupled geometric and texture PDE-based segmentation*. IEEE International Conference on Image Processing, vol. 2, pages 650–653, 2005.
- [Soille 1999] P. Soille. *Morphological image analysis*. Springer-Verlag Berlin, Heidelberg, 1999.
- [Stehman 1997] S. V. Stehman. *Selecting and interpreting measures of thematic classification accuracy*. Remote Sensing of Environment, vol. 62, pages 77–89, 1997.
- [Stein 1995] F. Stein et G. Medioni. *Map-based localization using the panoramic horizon*. In International Conference on Robotics and Automation (ICRA), 1995.
- [Strehl 2000] A. Strehl, J. Ghosh et R. Mooney. *Impact of similarity measures on web-page clustering*. The 17th National Conference on Artificial Intelligence : Workshop of Artificial Intelligence for Web Search (AAAI), pages 58–64, 2000.
- [Tamura 1978] H. Tamura, S. Mori et Y. Yamawaki. *Textural features corresponding to visual perception*. IEEE Transactions on Systems, Man, and Cybernetics, vol. 8, pages 460–473, 1978.
- [Tan 1994] T.S.C. Tan et J. Kittler. *Colour texture analysis using colour histogram*. IEEE Transactions on Image Processing Proceedings of Vision, Image and Signal Processing, vol. 141, pages 403–412, 1994.
- [Tardif 2008] J. Tardif, Y. Pavlidis et K. Daniilidis. *Monocular visual odometry in urban environments using an omnidirectional camera*. In IROS, 2008.
- [Tesar 2007] L. Tesar, D. Smutek, A. Shimizu et H. Kobatake. *Medical image segmentation using cooccurrence matrix based texture features calculated on weighted region*. ACST'07 Proceedings of the third conference on IASTED International Conference : Advances in Computer Science and Technology, pages 243–48, 2007.
- [Tuceryan 1998] M. Tuceryan et A. K. Jain. *Texture analysis*. In The Handbook of Pattern Recognition and Computer Vision, pages 207–248. Editions World Scientific Publishing Co, 1998.
- [Tupin 2008] F. Tupin et M. Sigelle. *Markov random fields*. ISTE Ltd and John Wiley & Sons, Digital Signal and Image Processing Series, 2008.
- [Turk 1991] M. Turk et A. Pentland. *Eigenfaces for recognition*. Journal of Cognitive Neuroscience, vol. 3, pages 71–86, 1991.
- [Ugarriza 2009] L.G. Ugarriza, E. Saber, S.R. Vantaram, V. Amuso, M. Shaw et R. Bhaskar. *Automatic image segmentation by dynamic region growth and multiresolution merging*. IEEE Transactions on Image Processing, vol. 18, pages 2275–2288, 2009.
- [Unnikrishnan 2005] R. Unnikrishnan et M. Hebert. *Measures of similarity*. Seventh IEEE Workshop on Computer Vision Applications, pages 394–400, 2005.
- [Unnikrishnan 2007] R. Unnikrishnan, C. Pantofaru et M. Hebert. *Toward objective evaluation of image segmentation algorithms*. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 29, pages 929–944, 2007.



- [Unser 1986] M. Unser. *Sum and difference histogram for texture classification*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, pages 118–125, 1986.
- [Usamentiaga 2006] R. Usamentiaga, D. F. Garcia, C. Lopez et D. Gonzalez. *A method for assessment of segmentation success considering uncertainty in the edge positions*. In EURASIP Journal on Applied Signal Processing, 2006.
- [Vandenbroucke 2000] N. Vandenbroucke. *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football*. PhD thesis, Université des sciences et technologies de Lille1, 2000.
- [Vapnik 1996] V. Vapnik. *The nature of statistical learning theory*. Springer, 1996.
- [Vese 2002] L. Vese et T. Chan. *A multiphase level set framework for image segmentation using the mumford and shah model*. In IJCV, 2002.
- [Viandier 2011] N. Viandier. *Modélisation et utilisation des erreurs de pseudodistances GNSS en environnement transport pour l'amélioration des performances de localisation*. PhD thesis, École centrale de Lille, 2011.
- [Vincent 1991] L. Vincent et P. Soille. *Watersheds in digital spaces : an efficient algorithm based on immersions simulations*. IEEE Trans. On Pattern Analysis and Machine Intelligence (PAMI), vol. 13, no. 16, pages 583–598, 1991.
- [Vincent 1993] L. Vincent. *Morphological Gray scale Reconstruction in Image Analysis : Applications and Efficient Algorithms*. IEEE Transactions on image processing, vol. 2, pages 176–201, 1993.
- [VisTex 1995] VisTex. *Vision texture database*. Massachusetts Institute of Technology, 1995.
- [Wang 2006] H. Wang, X.H. Wang, Y. Zhou et J. Yang. *Colour texture segmentation using quaternion-Gabor filters*. IEEE International Conference on Image Processing, pages 745–748, 2006.
- [Wirth 2006] M. Wirth, M. Fraschini, M. Masek et M. Bruynooghe. *Special Issue on Performance Evaluation in Image Processing*. In EURASIP Journal on Applied Signal Processing. Hindawi Publishing Corporation, 2006.
- [Xia 2006] Y. Xia, D. Feng et R. Zhao. *Adaptive segmentation of textured images by using the coupled Markov random field model*. IEEE Transactions on Image Processing, vol. 15, pages 3559–3566, 2006.
- [Yang 1995] L. Yang, F. Albregtsen, T. Lonnestad et P. Grøttum. *A supervised approach to the evaluation of image segmentation methods*. In Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns, pages 759–765, 1995.
- [Yang 2008] A.Y. Yang, J. Wright, Y. Ma et S. Sastry. *Unsupervised segmentation of natural images via lossy data compression*. Computer Vision and Image Understanding, vol. 110, pages 212–225, 2008.
- [Yanga 2008] X. Yanga, H. Rena et B. Lib. *Embedded zerotree wavelets coding based on adaptive fuzzy clustering for image compression*. Image and Vision Computing, vol. 26, pages 812–819, 2008.
- [Yasnoff 1979] W.A Yasnoff, W. Galbraith et J. W. Bacus. *Error measures for objective assessment of scene segmentation algorithms*. AQC, vol. 1, pages 107–121, 1979.

- [Yu 2008] S.Y. Yu, Y.Zhang, Y.G. Wang et J. Yang. *Unsupervised colour–texture image segmentation*. Journal of Shanghai Jiaotong University, vol. 13 E (1), pages 71–75, 2008.
- [Zeng 2011] W. Zeng et Y. Zhang. *A novel improvement to PCA for image classification*. International conference on Computer science and service system (CSSS), pages 1964–1967, 2011.
- [Zhang 1996] Y. J. Zhang. *A survey on evaluation methods for image segmentation*. Pattern Recognition, vol. 29, pages 1335–1346, 1996.
- [Zhang 2004] H. Zhang, J. E. Fritts et S. A. Goldman. *An entropybased objective evaluation method for image segmentation*. In ISET/SPIE's 16th Annual Symposium on Electronic Imaging Conference on Storage and Retrieval Methods and Applications for Multimedia, 2004.
- [Zhang 2006] W. Zhang et J. Kosecka. *Image based localization in urban environments*. In 3D Data Processing Visualization and Transmission (3DPVT), 2006.
- [Zhang 2008] H. Zhang, J. Fritts et S. Goldman. *Image segmentation evaluation : a survey of unsupervised methods*. Computer Vision and Image Understanding, vol. 110, pages 260–280, 2008.
- [Zheng 2006] Y. Zheng, J. Yang, Y. Zhou et Y. Wang. *Colour–texture based unsupervised segmentation using JSEG with fuzzy connectedness*. Journal of Systems Engineering and Electronics, vol. 17, pages 213–219, 2006.
- [Zheng 2007] C. Zheng, D.W. Sun et L. Zheng. *A new region-primitive method for classification of colour meat image texture based on size, orientation and contrast*. Meat Science, vol. 76, pages 620–627, 2007.
- [Zhu 1996] S. C. Zhu et A. Yuille. *Region competition : unifying snakes, region growing, and bayes/MDL for multiband image segmentation*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 18, pages 884–900, 1996.
- [Zoller 2002] T. Zoller, L. Hermes et J.M. Buhmann. *Combined colour and texture segmentation by parametric distributional clustering*. Proceedings of the 16 th International Conference on Pattern Recognition (ICPR), vol. 2, pages 627–630, 2002.



# TABLE DES FIGURES

1.1	Sensibilité spectrale des trois types de cônes photorécepteurs couleur de l'œil humain et des bâtonnets (source : <a href="http://tpelavision.chez-alice.fr/oeil.htm">http://tpelavision.chez-alice.fr/oeil.htm</a> ). . . . .	21
1.2	Représentation 3D de deux images couleur dans un espace muni du repère $(C_1, C_2, C_3)$ à l'aide du logiciel "Color Inspector 3D" (source : <a href="http://www.f4.fhtw-berlin.de/~barthel/ImageJ/ColorInspector//help.htm">http://www.f4.fhtw-berlin.de/~barthel/ImageJ/ColorInspector//help.htm</a> ). . . . .	22
1.3	Exemples de textures répétitives prises de la base de texture VISTEX [VisTex 1995]. . . . .	22
1.4	Différents types de textures extraites de la base de texture VISTEX [VisTex 1995]. . . . .	23
1.5	Différentes catégories de texture. . . . .	23
1.6	Les différentes classes d'attributs de texture de la littérature. . . . .	24
1.7	Cliques. . . . .	26
1.8	Exemples de voisinages utilisés pour le calcul des LBP : à gauche $N = 8$ , au centre $N = 12$ et à droite $N = 16$ . . . . .	27
1.9	Illustration du principe de calcul du LBP. . . . .	28
1.10	Illustration des différents histogrammes mono-dimensionnels et tridimensionnel d'une image couleur dans l'espace couleur $(R, G, B)$ (de haut en bas et de gauche à droite : image originale couleur, histogramme tridimensionnel obtenue par le logiciel <b>Color Inspector 3D</b> et histogrammes mono-dimensionnels de la première, deuxième et troisième composante couleur). . . . .	30
1.11	Un vecteur de translation $\vec{R}$ en fonction des paramètres $\theta$ et $\ \vec{R}\ $ . . . . .	31
1.12	Exemple de différents vecteurs de translation $\vec{R}$ . . . . .	32
1.13	Taux (en noir) et direction (en violet) de croissance des travaux publiés sur les méthodes de segmentation couleur/texte entre les années 1984 et 2008 [Ilea 2011]. . . . .	34
1.14	Les différents types des méthodes de segmentation utilisant conjointement la couleur et la texture. . . . .	35
1.15	Construction d'un vecteur d'attributs couleur/texte. Une fenêtre de taille $w \times w$ de chaque composante couleur de l'espace CIE Lab est filtrée par un filtre gaussien. Les valeurs obtenues pour chaque composante couleur sont stockées dans un seul vecteur [Yang 2008]. . . . .	36
1.16	Principe de l'algorithme Jseg. . . . .	41

1.17	Technique de génération d'un descripteur cubique : Pour un pixel $p$ d'une image dans l'espace couleur $(R, G, B)$ et dans un voisinage de taille $N_w \times N_w$ ( $N_w = 7$ ), l'histogramme cubique de taille $N_b = q^3$ ( $q = 5$ ) est obtenu en incrémentant de (+1) la valeur de la case correspondant aux coordonnées $R_{p'}, V_{p'}$ et $B_{p'}$ associées au pixel $p'$ appartenant au voisinage de $p$ [Mignotte 2008]. . . . .	42
1.18	Le gradient orienté des histogrammes [Arbelaez 2011]. De gauche à droite : image d'intensité avec un disque circulaire centré en un pixel donné, gradient résultant avec un disque de rayon 5 pixels et une orientation $\theta = \frac{\pi}{4}$ . . . . .	49
1.19	Les 17 filtres gaussiens [Arbelaez 2011]. . . . .	50
1.20	Formation du gradient composite par fusion des gradients couleur et texture [Cohen 2010a]. . . . .	53
1.21	Résultats des principales méthodes de segmentation sur une sélection d'images de la base de BERKLEY [Berkeley] (de haut en bas : images originales, méthode de segmentation CTM [Yang 2008], méthode de segmentation MeanShift [Comaniciu 2002], méthode de segmentation Jseg [Deng 2001] et méthode de segmentation proposée par FELZENSZ WALB et HUTTENLOCHER [Felzenszwalb 2004]). . . . .	55
1.22	Résultats des principales méthodes de segmentation sur une sélection d'images de la base de BERKLEY [Berkeley] (de haut en bas : images originales, méthode de segmentation proposée par CHAN et VESE [Chan 2001], méthode de segmentation par combinaison couleur, texture et luminance [Arbelaez 2011], méthode de segmentation par combinaison couleur/texture fixe [Angulo 2007] et méthode de segmentation par combinaison couleur/texture adaptative [Cohen 2009]). . . . .	56
2.1	Quelle segmentation et quelle métrique d'évaluation choisir ? . . . . .	58
2.2	Classification des méthodes d'évaluation proposée par ZHANG [Zhang 1996]. . . . .	59
2.3	Classification des méthodes d'évaluation proposée par JIANG et al. [Jiang 2006]. . . . .	59
2.4	Classification des méthodes d'évaluation proposée par MCGUINNESS [McGuinness 2009]. . . . .	60
2.5	Illustration de plusieurs segmentations de référence réalisées pour une même image. . . . .	62
2.6	La matrice de confusion définie pour deux classifications $S$ et $VT$ . . . . .	63
2.7	Segmentation automatique par l'algorithme MeanShift et deux segmentations manuelles, d'une image de la base de BERKELEY. . . . .	70
2.8	Échantillons d'images issues de différentes catégories de la base de MCGILL. . . . .	75
2.9	Une sélection d'images avec quelques-unes de leurs segmentations de référence extraites de la base de BERKELEY. . . . .	76
2.10	Une sélection d'images de la base de PASCAL VOC. . . . .	77

2.11 Une sélection de textures uniformes et trois scènes naturelles extraites de la base de texture VISTEX. . . . .	77
2.12 Une sélection d'images de la base de texture OUTEX. . . . .	78
2.13 Une sélection de textures mono-spectrales et multi-spectrales de la base de PRAGUE. . . . .	78
2.14 Différentes textures extraites de la librairie de texture. . . . .	79
2.15 Exemple de variation de la performance de la segmentation sur une image en fonction de cinq vérités-terrain. . . . .	80
2.16 Processus de l'évaluation par les mesures de VINET. . . . .	81
2.17 Évaluation moyenne (avec les métriques de RAND $\mathcal{PR}$ , de JACCARD $\mathcal{J}$ et de FOWLKES et MALLOWS $\mathcal{F}$ sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison $\alpha$ . . . . .	84
2.18 Évaluation moyenne (avec les métriques de Bipartite Graph Matching $\mathcal{BGM}$ , d'information mutuelle normalisée $\mathcal{NMI}$ et de variation d'information $\mathcal{VI}$ sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison $\alpha$ . . . . .	85
2.19 Évaluation moyenne (avec les métriques de MARTIN $\mathcal{LCE}$ , $\mathcal{GCE}$ et de VINET $\mathcal{V}$ sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison $\alpha$ . . . . .	86
2.20 Évaluation moyenne (avec les métriques d'Erreur quadratique moyenne $\mathcal{MSE}$ , de ZEBODJ $\mathcal{Z}$ et de BORSOTTI $\mathcal{B}$ sur 200 images de la base de BERKELEY) de la segmentation par combinaison fixe couleur/texture en fonction du paramètre de combinaison $\alpha$ . . . . .	87
2.21 Évaluation (sur 33 images de la base de VisTex) des méthodes de segmentation couleur/texture pour chaque image et en moyenne. . . . .	95
2.22 Illustration des segmentations sur huit images de la base de BERKELEY [Arbelaez 2011] (de haut en bas : les images originales, images segmentées par CTB, images segmentées par MeanShift, images segmentées par FCR, images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, images segmentées par CTM et images segmentées par Jseg). . . . .	97
2.23 Illustration des segmentations sur huit images de la base de BERKELEY [Arbelaez 2011] (de haut en bas : les images originales, images segmentées par Ncuts, images segmentées par CHAN-VESE, images segmentées par combinaison adaptative, images segmentées par combinaison supremum et images segmentées par combinaison fixe). . . . .	98
2.24 Illustration des segmentations sur cinq images de la base de texture VisTex [VisTex 1995](de haut en bas : les images originales, images segmentées par MeanShift, images segmentées par FCR, images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, images segmentées par CTM, images segmentées par Jseg et images segmentées par Ncuts). . . . .	99



2.25	Illustration des segmentations sur cinq images de la base de texture VisTex [VisTex 1995] (de haut en bas : les images originales, images segmentées par CHAN-VESE, images segmentées par combinaison adaptative, images segmentées par combinaison supremum et images segmentées par combinaison fixe). . . . .	100
3.1	Illustration du gradient morphologique couleur calculé sur une image de la base de BERKELEY. . . . .	103
3.2	Illustration du gradient de DI-ZENZO sur une image couleur de la base de BERKELEY. . . . .	105
3.3	Illustration du descripteur de texture basé sur les filtres de Gabor sur une image couleur issue de la base de BERKELEY dans l'espace couleur $(R, G, B)$ . . . . .	106
3.4	Illustration des matrices de co-occurrences de textures de gel calculées sur une fenêtre de $60 \times 60$ pixels sur 16 niveaux de gris ( $N = 16$ ) [Hibon 2010]. . . . .	107
3.5	Gradients calculés à partir d'attributs d'HARALICK. . . . .	111
3.6	Illustration d'une simplification d'une image issue de la base de BERKELEY par un filtre ASF avec différents ordres. . . . .	112
3.7	Illustration d'un filtrage de nivellement sur une image de la base de BERKELEY bruitée avec un bruit gaussien (de moyenne 5 et d'écart type 10) et avec trois images marqueurs (obtenues avec un filtre gaussien, un filtre moyennneur et un filtre ASF). . . . .	113
3.8	Calcul de la couche de texture. . . . .	114
3.9	Analyse multi-échelle locale pour le calcul du gradient morphologique de texture $\mathcal{G}_{ASF}$ basé sur le filtrage par ASF. . . . .	116
3.10	Synopsis global de la combinaison couleur-texture basée sur l'ACP. . . . .	119
3.11	Exemple de trois dispersions géométriques possibles d'un nuage de points (NP) 3D. . . . .	122
3.12	Bassins versants et ligne de partage des eaux [Vincent 1991]. . . . .	124
3.13	Exemples de réductions de germes avec le résultat de segmentation correspondant. . . . .	125
3.14	Synopsis global de la méthode de segmentation proposée. . . . .	126
3.15	Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par CTB, les images segmentées par MeanShift, les images segmentées par FCR et les images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER). . . . .	133
3.16	Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par Ncuts, les images segmentées par CHAN-VESE, les images segmentées par CTM et les images segmentées par Jseg). . . . .	134

3.17	Illustration des résultats de segmentation sur cinq images de la base de BERKELEY (de haut en bas : les images originales, les images segmentées par combinaison adaptative, les images segmentées par combinaison supremum, les images segmentées par combinaison fixe avec le paramètre de pondération $\alpha$ optimal associé à chaque image et les images segmentées par combinaison par Analyse en Composantes Principales. . . . .	135
3.18	Illustration des résultats de segmentation sur quatre images de la base de VISTEX (de haut en bas : les images originales, les images segmentées par MeanShift, les images segmentées par FCR, les images segmentées par la méthode de FELZENSZWALB et HUTTENLOCHER, les images segmentées par CTM, les images segmentées par Jseg et les images segmentées par Ncuts). . . . .	138
3.19	Illustration des résultats de segmentation sur quatre images de la base de VISTEX (de haut en bas : les images originales, les images segmentées par la segmentation de CHAN-VESE, les images segmentées par combinaison adaptative, les images segmentées par combinaison supremum, les images segmentées par combinaison fixe avec le paramètre de pondération $\alpha$ optimal associé à chaque image et les images segmentées par combinaison par comparaison par Analyse en Composantes Principales. . . . .	139
4.1	Illustration du phénomène de multi-trajets en milieu urbain. . . . .	142
4.2	Véhicules instrumentés de l'UTBM/IRTES-SET (à gauche) et de l'IFSTTAR-LEOST (à droite). . . . .	145
4.3	Illustration de cinq images issues de la base du projet LOCOPROL. . . . .	146
4.4	Illustration d'un trajet d'une compagne d'acquisition réalisée à Belfort, dans le cadre du projet CAPLOC avec 5 des 149 images acquises en mobilité. . . . .	146
4.5	Récapitulatif des étapes de la stratégie de caractérisation de l'environnement de réception des signaux GNSS. . . . .	148
4.6	Résultats de segmentation et de classification obtenue avec l'algorithme Fisher pour chaque combinaison couleur/texte sur une sélection d'images de la base du projet LOCOPROL (de haut en bas : images originales, vérités-terrain, images segmentées puis classées par la combinaison fixe, la combinaison hybride et par la combinaison proposée). . . . .	151
4.7	Résultats de segmentations et de classifications obtenues avec l'algorithme Fisher pour chaque combinaison couleur/texte sur une sélection d'images de la base du projet CAPLOC (de haut en bas : images originales, vérités-terrain, images segmentées puis classées par la combinaison fixe, par la combinaison hybride et par la combinaison proposée). . . . .	153
4.8	Illustration de la présence de nuages et de forte illumination du soleil sur une sélection d'images issues de la base CAPLOC. . . . .	154
4.9	Les différents types de filtre dans la littérature. . . . .	155

4.10 Illustration du principe du filtre médian avec $W = 3$ . . . . .	156
4.11 Illustration du lissage par filtre médian avec différentes valeurs de voisinage $W$ . . . . .	156
4.12 Illustration du lissage par filtre moyenneur avec différentes valeurs de voisinage $W$ . . . . .	157
4.13 Illustration du lissage par filtre gaussien avec différentes valeurs de $\sigma$ . . . . .	157
4.14 Illustration du lissage par filtre exponentiel avec différentes valeurs de $\alpha$ . . . . .	158
4.15 Illustration de l'évolution du lissage par filtre de MALIK-PERONA en fonction du nombre d'itérations $itr$ . . . . .	158
4.16 Les 9 domaines de NAGAO. . . . .	159
4.17 Illustration du lissage par filtre de NAGAO. . . . .	159
4.18 Principe du filtre adaptatif basé sur le voisinage. . . . .	160
4.19 Illustration du lissage par filtre adaptatif basé sur le voisinage avec différentes valeur du seuil. . . . .	160
4.20 Principe de calcul du filtre SNN. . . . .	160
4.21 Illustration du lissage par filtre adaptatif Symetric Nearest Neighbourhood. . . . .	161
4.22 Effet de la RGD sur une fonction 1D. . . . .	163
4.23 Illustration de l'effet de la simplification par RGD sur une image issue de la base CAPLOC, avec la représentation 3D de l'intensité. . . . .	164
4.24 Influence de la valeur du paramètre de contraste $H$ sur la simplification d'image par RGD. . . . .	164
4.25 La classification supervisée. . . . .	169
4.26 Résultats de classification " ciel " / " non-ciel " avec différents algorithmes non-supervisés en fonction de la méthode de simplification utilisée sur la base LOCOPROL. . . . .	172
4.27 Résultats de classification " ciel " / " non-ciel " avec différents algorithmes supervisés en fonction de la méthode de simplification utilisée sur la base LOCOPROL. . . . .	173
4.28 Résultats de classification " ciel " (en bleu) / " non ciel " (en marron) sur 5 images de la base d'images du projet LOCOPROL avec différents classifieurs (de haut en bas : Fisher (composante couleur bleue), KMLocal (version Hybrid), FuzzyCmeans ( $m=2$ ), SRM, Bayes, Kppv ( $nbKppv=6$ ), SVM). . . . .	175
4.29 Résultats de classification ciel/non-ciel avec différents algorithmes non-supervisés en fonction de la méthode de simplification utilisée sur la base CAPLOC. . . . .	176
4.30 Résultats de classification ciel/non-ciel avec différents algorithmes supervisés en fonction de la méthode de simplification utilisée sur la base CAPLOC. . . . .	177

4.31	Résultats de classification " ciel" (en bleu) / " non ciel " (en marron) sur 5 images de la base d'images du projet CAPLOC avec différents algorithmes de classification (de haut en bas : Fisher (composante couleur bleue), KMlocal (version Loyds), FuzzyCmeans (m=2), SRM, Bayes, Kppv (nbKppv=6), SVM).	179
5.1	Exemple d'une bonne et d'une mauvaise configuration géométrique des satellites.	185
5.2	Signal reçu à partir d'un satellite NLOS, après une réflexion.	186
5.3	Synopsis de la stratégie globale pour l'amélioration de la précision lors du calcul de la position d'un récepteur GNSS.	187
5.4	Dispositif de calibrage [Nowak 2005].	189
5.5	Illustration de la détection du ciel avec repositionnement des satellites dans l'image (de haut en bas : images initiales, images classées par l'algorithme Fisher après filtrage médian).	189
5.6	Principe de la détection de la ligne d'horizon.	190
5.7	Détection d'une ligne d'horizon.	190
5.8	Trajectoire étudiée : positions de référence (en rouge) et positions estimées (en bleue).	192
5.9	Trajectoires obtenues en n'utilisant que les satellites LOS pour le calcul des positions.	193
5.10	Illustration des différences entre les deux cas d'utilisation.	194
5.11	L'évolution dans le temps du nombre total de satellites et du nombre de satellites LOS.	194
5.12	L'évolution dans le temps du GDOP des satellites LOS (cas 1 et 2) et de tous les satellites.	195
5.13	Illustration de la répartition géométrique des satellites sur les deux parties du trajet.	196
5.14	Comparaison entre les trois solutions de calcul de position : la trajectoire rouge est la référence (GPS-RTK), la trajectoire bleue issue de la solution utilise tous les satellites, la trajectoire verte est issue de la solution qui n'utilise que les satellites LOS, la trajectoire violette est issue de la solution par combinaison (exclusion des satellites NLOS si la valeur du GDOP est inférieur au seuil choisi ( $Seuil_{GDOP} = 3$ )).	197
5.15	Vue aérienne (Google Earth) des trajectoires obtenues avec les trois solutions (en rouge : la référence (GPS-RTK), en bleu : tous les satellites sont utilisés, en vert : les satellites LOS sont uniquement utilisés, en violet : la solution par combinaison).	198
5.16	Exemple, sur deux images issues de la base du projet CAPLOC, de changement de l'état de satellites de NLOS à LOS, en fonction de la classification de la végétation en masque ou non.	199

5.17 Comparaison entre les deux hypothèses ( $LOS \equiv Ciel$ et $LOS \equiv Ciel-Végétation$ ). . . . .	200
5.18 Principe de la projection orthogonale pour le calcul de l'erreur de position [Viandier 2011]. . . . .	201
5.19 Courbes d'erreur en fonction de la position estimée avec chacune des trois solutions d'estimation de position. . . . .	202
5.20 Fonction de répartition de l'erreur de positionnement selon la solution utilisée.	202

# LISTE DES TABLES

1.1	Avantages et limites des principales méthodes de segmentation d'images de la littérature utilisant conjointement les informations de couleur et de texture. . . . .	54
2.1	Évaluation, avec différentes métriques, de la segmentation illustrée dans la figure 2.7. . . . .	70
2.2	Désignation des différentes combinaisons couleur/texture testées. . . . .	82
2.3	Performances moyennes (calculées sur 200 images de la base de BERKELEY avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture) de la segmentation par combinaison couleur/texture fixe [Angulo 2007] ( $\alpha^*$ correspond à la valeur optimale du paramètre de combinaison $\alpha$ ). . . . .	83
2.4	Résultats (avec une évaluation par $MSE$ ) de la segmentation par combinaison fixe couleur/texture sur cinq images de la base de BERKELEY ( $\alpha^*$ correspond à la valeur optimale du paramètre de combinaison $\alpha$ ). . . .	88
2.5	Performances moyennes (calculées sur 200 images de la base de BERKELEY avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture) des segmentations par combinaison adaptative et par combinaison supremum [Cohen 2009]. . . . .	89
2.6	Performances moyennes des principales méthodes de segmentation de la littérature sur 200 images de la base de BERKELEY : MS MeanShift, JSEG, FH Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR, CTM Compression-based Texture Merging, CV segmentation de CHAN-VESE, CTB segmentation basée sur la combinaison de la couleur, la texture et la luminosité (Color, Texture and Brightness), CF segmentation basée sur la combinaison fixe, CA segmentation basée sur la combinaison adaptative et CS segmentation basée sur la combinaison supremum. . . .	91
2.7	Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec les métriques d'évaluation non-supervisée $MSE$ , $\mathcal{Z}$ de ZEBOU DJ et $\mathcal{B}$ de BORSOTTI) de la méthode de segmentation basée sur une combinaison fixe couleur/texture en fonction des gradients couleur/texture et du paramètre de pondération $\alpha$ . . . . .	92
2.8	Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec $MSE$ , $\mathcal{Z}$ et $\mathcal{B}$ ) de la méthode de segmentation basée sur la combinaison fixe couleur/texture en fonction de la valeur optimale $\alpha^*$ du paramètre de combinaison. . . . .	93



2.9 Performances moyennes (calculées sur 33 images de la base de texture VISTEX avec $MSE$ , $Z$ et $B$ ) des méthodes de segmentation basées sur les combinaisons adaptative et supremum des informations de couleur et de texture. . . . .	93
2.10 Comparatif (sur 33 images de texture de la base de VisTex) des performances moyennes des principales méthodes de segmentation de la littérature . . . . .	94
3.1 Performances moyennes (calculées sur 200 images de la base de BERKELEY) avec les différentes métriques d'évaluation et en fonction des couples gradients couleur/texture de la méthode de segmentation par la combinaison couleur/texture proposée. . . . .	128
3.2 Comparaison, sur 200 images de la base de BERKELEY, des résultats de la segmentation par la combinaison couleur/texture proposée et des principales méthodes de segmentation de la littérature : (MS = MeanShift, JSEG, FH = Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR = Fusion Clustering Results, CTM = Compression-based Texture Merging, CV = segmentation de CHAN-VESE, CTB segmentation basée sur la combinaison de la couleur, la texture et la luminosité (Color, Texture and Brightness), CF = segmentation basée sur la combinaison fixe, CA = segmentation basée sur la combinaison adaptative, CS = segmentation basée sur la combinaison supremum et C-ACP = segmentation basée sur la combinaison par Analyse en Composantes Principales). Les valeurs en rouge (respectivement en bleue) désignent la première meilleure (respectivement la deuxième meilleure) performance. . . . .	132
3.3 Performance de la méthode de segmentation basée sur la combinaison couleur/texture proposée sur la base de texture VISTEX, en fonction des différents couples des gradients couleur et texture. . . . .	136
3.4 Comparaison, sur 33 images de la base de VISTEX, des résultats de segmentation basée sur la combinaison couleur/texture proposée et les principales méthodes de segmentation de la littérature : (MS = MeanShift, JSEG, FH = Segmentation de FELZENSZWALB et HUTTENLOCHER, FCR = Fusion Clustering Results, CTM = Compression-based Texture Merging, CV = segmentation de CHAN-VESE, CF = segmentation basée sur la combinaison fixe, CA = segmentation basée sur la combinaison adaptative, CS = segmentation basée sur la combinaison supremum et C-ACP = segmentation basée sur la combinaison par Analyse en Composantes Principales). Les valeurs en rouge (respectivement en bleue) désignent la première meilleure (respectivement la deuxième meilleure) performance. . . . .	137
4.1 Résultats de classification binaire (" ciel "/" non ciel ") en fonction de la méthode de combinaison couleur/texture utilisée (fixe, hybride et proposée) et des couples gradients couleur/texture, sur la base d'images du projet LOCOPROL. . . . .	150

4.2	Résultats de classification binaire (" ciel "/" non ciel ") en fonction de la méthode de combinaison couleur/texture utilisée (fixe, hybride et proposée) et des couples gradients couleur/texture, sur la base d'images du projet CAPLOC. . . . .	152
4.3	Résultats de classification " ciel " / " non ciel " avec les différents algorithmes de classification, méthodes de simplification utilisés et paramètres, sur la base d'images du projet LOCOPROL. . . . .	174
4.4	Résultats de classification " ciel " / " non ciel " avec différents algorithmes de classification, méthodes de simplification utilisés et paramètres, sur la base d'images du projet CAPLOC. . . . .	178
4.5	Temps de traitement de l'approche en fonction du couple méthode de simplification / algorithme de classification. . . . .	178
5.1	Performances des différentes approches de calcul de position proposées. .	203





## Résumé :

En segmentation d'images, les informations de couleur et de texture sont très utilisées. Le premier apport de cette thèse se situe au niveau de l'utilisation conjointe de ces deux sources d'informations. Nous proposons alors une méthode de combinaison couleur/texture, adaptative et non paramétrique, qui consiste à combiner un (ou plus) gradient couleur et un (ou plus) gradient texture pour ensuite générer un gradient structurel utilisé comme image de potentiel dans l'algorithme de croissance de régions par LPE. L'originalité de notre méthode réside dans l'étude de la dispersion d'un nuage de point 3D dans l'espace, en utilisant une étude comparative des valeurs propres obtenues par une analyse des composantes principales de la matrice de covariance de ce nuage de points. L'approche de combinaison couleur/texture proposée est d'abord testée sur deux bases d'images, à savoir la base générique d'images couleur de BERKELEY et la base d'images de texture VISTEX. Cette thèse s'inscrivant dans le cadre des projets ViLoc (RFC) et CAPLOC (PREDIT), le deuxième apport de celle-ci se situe au niveau de la caractérisation de l'environnement de réception des signaux GNSS pour améliorer le calcul de la position d'un mobile en milieu urbain. Dans ce cadre, nous proposons d'exclure certains satellites (NLOS dont les signaux sont reçus par réflexion voir totalement bloqués par les obstacles environnants) dans le calcul de la position d'un mobile. Deux approches de caractérisation, basées sur le traitement d'images, sont alors proposées. La première approche consiste à appliquer la méthode de combinaison couleur/texture proposée sur deux bases d'images réelles acquises en mobilité, à l'aide d'une caméra fisheye installée sur le toit du véhicule de laboratoire, suivie d'une classification binaire permettant d'obtenir les deux classes d'intérêt « ciel » (signaux LOS) et « non ciel » (signaux NLOS). Afin de satisfaire la contrainte temps réel exigée par le projet CAPLOC, nous avons proposé une deuxième approche basée sur une simplification de l'image couplée à une classification pixellaire adaptée. Le principe d'exclusion des satellites NLOS permet d'améliorer la précision de la position estimée, mais uniquement lorsque les satellites LOS (dont les signaux sont reçus de manière direct) sont géométriquement bien distribués dans l'espace. Dans le but de prendre en compte cette connaissance relative à la distribution des satellites, et par conséquent, améliorer la précision de localisation, nous avons proposé une nouvelle stratégie pour l'estimation de position, basée sur l'exclusion des satellites NLOS (identifiés par le traitement d'images), conditionnée par l'information DOP, contenue dans les trames GPS.

**Mots-clés :** Segmentation d'images; Classification de pixels; Couleur; Texture; Combinaison couleur/texture; ACP; Localisation; GNSS; Transport.

## Abstract:

Color and texture are two main information used in image segmentation. The first contribution of this thesis focuses on the joint use of color and texture information by developing a robust and non parametric method combining color and texture gradients. The proposed color/texture combination allows defining a structural gradient that is used as potential image in watershed algorithm. The originality of the proposed method consists in studying a 3D points cloud generated by color and texture descriptors, followed by an eigenvalue analysis. The color/texture combination method is firstly tested and compared with well known methods in the literature, using two databases (generic BERKELEY database of color images and the VISTEX database of texture images). The applied part of the thesis is within ViLoc project (funded by RFC regional council) and CAPLOC project (funded by PREDIT). In this framework, the second contribution of the thesis concerns the characterization of the environment of GNSS signals reception. In this part, we aim to improve estimated position of a mobile in urban environment by excluding NLOS satellites (for which the signal is masked or received after reflections on obstacles surrounding the antenna environment). For that, we propose two approaches to characterize the environment of GNSS signals reception using image processing. The first one consists in applying the proposed color/texture combination on images acquired in mobility with a fisheye camera located on the roof of a vehicle and oriented toward the sky. The segmentation step is followed by a binary classification to extract two classes « sky » (LOS signals) and « not sky » (NLOS signals). The second approach is proposed in order to satisfy the real-time constraint required by the application. This approach is based on image simplification and adaptive pixel classification. The NLOS satellites exclusion principle is interesting, in terms of improving precision of position, when the LOS satellites (for which the signals are received directly) are well geometrically distributed in space. To take into account the knowledge of satellite distribution and then increase the precision of position, we propose a new strategy of position estimation, based on the exclusion of NLOS satellites (identified by the image processing step), conditioned by DOP information, which is provided by GPS data.

**Keywords:** Image segmentation; Classification of pixels; Color; Texture; Colour/Texture Combination; PCA; Localisation; GNSS; Transport.

SPIM